

Basic entanglement distillation with realistic noise

Vikesh Siddhu , Erick Winston , David C. McKay *, and Ali Javadi-Abhari [†]
IBM Quantum, IBM T.J. Watson Research Center, Yorktown Heights, New York 10598, USA



(Received 25 April 2025; accepted 27 August 2025; published 17 October 2025)

Entanglement distillation is a key component of modular quantum computing and long-range quantum communications. However, this powerful tool to reduce noise in entangled states is difficult to realize in practice for two main reasons. First, operations used to carry out distillation inject noise they seek to remove. Second, the extent to which distillation can work under realistic device noise is less well studied. In this work, we both simulate distillation using a variety of device noise models and perform distillation experiments on fixed-frequency IBM devices. We find reasonable agreement between experimental data and simulation done using Pauli and non-Pauli noise models. In our data, we find broad improvement when the metric of success for distillation is to improve average Bell fidelity under effective global depolarizing noise, remove coherent errors, or improve the Bell fidelity of mildly degraded Bell pairs. We pave the way to obtain broad improvement from distillation under a stricter, but practically relevant, metric: distill (physically) nonlocal Bell pairs with higher fidelity than possible to obtain with other available methods. Our results also help understand metrics and requirements for quantum devices to use entanglement distillation as a primitive for modular computing.

DOI: [10.1103/xx9g-pkht](https://doi.org/10.1103/xx9g-pkht)

I. INTRODUCTION

Entanglement is an important nonclassical resource used in quantum computation and communication. However, this resource depletes due to noise that inevitably affects physical systems. This noise may be removed using entanglement distillation (also called purification) [1,2]. At a high level, entanglement distillation entangles many copies of lower-fidelity Bell states such that we can postselect on a subspace to obtain higher-fidelity Bell pairs. Therefore, this basic technique generally requires a high throughput of Bell pairs and its simplest variants are nondeterministic. However, via enough rounds of distillation and consumption of Bell pairs, one can hope to achieve high fidelities [3] (see also recent work [4]). In the context of quantum error correction (QEC) [5,6] and quantum compilation [7], high-fidelity Bell pairs can enable measurement and (gate) teleportation. When passing states through certain channels, distillation can even outperform QEC for communication [2].

There is a vast body of work dedicated to entanglement distillation [8–23]. From a theoretical perspective, understanding the best rates for distilling entanglement and methods for achieving these rates in practice is an important area of study. In practice, entanglement distillation plays a central role in both long-distance quantum communication [3,24,25] and scaling quantum computation, for instance using modular

architectures [26,27]. A variety of studies aim at optimizing protocols in these settings [28–32].

Due to its importance, a variety of prior experiments on various physical platforms including optical [33–37], trapped ion [38], solid state [39], and superconducting [40] qubits have reported entanglement distillation. Such proof-of-concept studies have opened the way for using entanglement distillation not only in quantum networks but also modular architectures for quantum computing. However, a variety of challenges exist in understanding the role distillation can play in practice, both in terms of finding appropriate distillation protocols that can be implemented and in terms of the efficacy of using distillation protocols over other strategies for distributing Bell pairs given the noise in the hardware components implementing these protocols.

Carrying out distillation on devices can be nontrivial and analyzing it requires care for noise sources different from those considered in a large part of entanglement distillation literature. For instance, a number of theoretical studies assume that the entangled states experience independent identically distributed (iid) noise (typically Pauli noise). Furthermore, many assume noiseless gates, measurements, and ancilla. In practice, entangled states experience non-iid noise. For instance, the natural way to create two physically nonlocal Bell pairs (a physically nonlocal Bell pair refers to a Bell state on physically unconnected qubits) may add noise differently to each pair [41]. This noise need not be Pauli noise [42]; for instance, most qubits have a non-negligible T_1 time, which is a non-Pauli error. The gates used to carry out entanglement distillation also add noise. Measurements done during or after distillation are also noisy and, typically, of long duration. These measurement errors not only affect the distillation protocol, but also the protocols used to certify how well distillation works. Finally, error-free ancillas, used in certain nested schemes for distillation, need not be available either.

*Contact author: dcmckay@us.ibm.com

[†]Contact author: ali.javadi@ibm.com

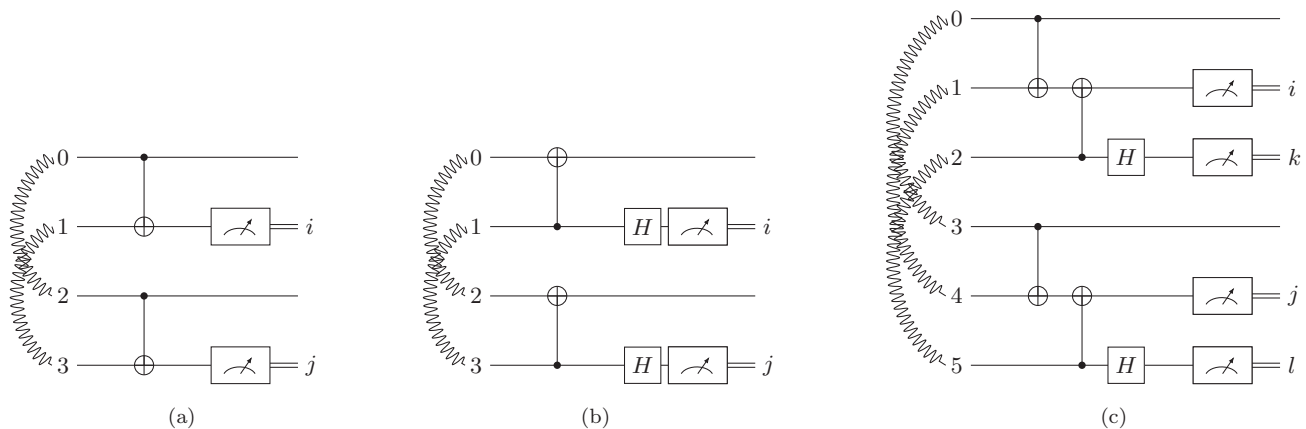


FIG. 1. Basic circuits for entanglement distillation. (a) The Z_{2B} recurrence protocol for entanglement distillation: Two-qubit states ρ_{02} and ρ_{13} are distilled as follows. First CNOT_{01} and CNOT_{23} gates are applied, next qubits 1 and 3 are measured in the Z basis, resulting in outcomes i and j , respectively. When $i = j$ the state on qubits 0 and 2 is retained, (b) The X_{2B} recurrence protocol for entanglement distillation: Two-qubit states ρ_{02} and ρ_{13} are distilled as follows. First CNOT_{10} and CNOT_{32} gates are applied, next qubits 1 and 3 are measured in the X basis, resulting in outcomes i and j , respectively. When $i = j$ the state on qubits 0 and 2 is retained, and (c) The ZX_{3B} protocol for entanglement distillation: Three two-qubit states ρ_{03} , ρ_{14} , ρ_{25} are distilled to a single two-qubit state. If outcomes of the $Z_0Z_1Z_2$ and $Z_3Z_4Z_5$ measurements and X_1X_2 and X_3X_4 measurements agree with one another, i.e., $i = j$ and $k = l$, respectively, the state on qubits 0 and 3 is retained.

In this work, we numerically and experimentally explore entanglement distillation on IBM's superconducting qubits. In our numerical exploration, we include noise on both the Bell pairs being distilled and the components (two-qubit gates and measurements) carrying out this distillation. For local Pauli noise, we report gate and measurement noise parameters for three different distillation protocols to improve the Bell fidelity of noisy Bell pairs (with possibly unequal initial Bell fidelity). For global depolarizing noise on the Bell pairs, we also report experimental results that we fit numerically. In further experimental exploration, we report the performance of three different distillation protocols as a function of the input Bell fidelity that we degrade by idling the Bell states before distillation. We model the waiting time error using a non-Pauli noise model (using the damping-dephasing channel [42,43]) and include ZZ crosstalk to obtain reasonable agreement with experimental data both when distillation provides an improvement in Bell fidelity and when it does not.

Our results show that (1) simple noise parameters of a device can be turned into a non-Pauli noise model to capture essential features of our entanglement distillation protocols; (2) using the noise model it is possible to pick one distillation protocol in favor of another; and (3) depending on the metric of success for distillation, one may or may not find broad improvement from the simplest distillation protocols.

The rest of this paper is organized as follows. Using standard notation (see Appendix A), we first summarize well-known ideas of entanglement distillation in Sec. II. These ideas include the recurrence protocol, what we call the $\{ZX_{3B}\}$ distillation protocol (also called double selection in Ref. [17]), impact of distillation on Bell pairs of unequal Bell fidelity, and distillation under global depolarizing noise. The next section, Sec. III A, is devoted to studying the effects of gate noise, measurement noise, and local depolarizing noise on Bell state preparation and entanglement distillation. In Sec. III B, we augment the noise model of the previous section by adding global depolarizing noise to qubits as they wait prior to being

distilled. We generate experimental data on superconducting qubits to mimic depolarizing noise using twirling. Finally, in Sec. IV we report results on the experimental study of three distillation protocols, two recurrence-type protocols, and the ZX_{3B} protocol, as a function of idling noise on superconducting qubits; we match these data numerically using a damping-dephasing (non-Pauli) noise model also described in that section.

II. ENTANGLEMENT DISTILLATION

In this section, we summarize the concept of distillation based on and extending the recurrence protocol [1,2] and show some simple calculations based on depolarizing models focusing on how the protocols suffer if there is an imbalance in the fidelities of the input states. Here, fidelity refers to what we call the Bell fidelity, $F := \langle \phi | \rho | \phi \rangle$, between a two-qubit density operator ρ and a fixed maximally entangled state $|\phi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$.

A. Recurrence (Z_{2B} , X_{2B})

The minimal recurrence protocols, using two Bell states (four qubits—0, 1, 2, and 3), are shown in Figs. 1(a) and 1(b). In Fig. 1(a), we show Z_{2B} recurrence where the circuit measures Z_0Z_1 and Z_2Z_3 , and postselects for the case where these two measurements, i and j , respectively, equal each other. Replacing these ZZ measurements with XX measurements results in what we call X_{2B} recurrence, shown in Fig. 1(b).

Due to the nature of the measurement, Z_{2B} can detect single X or Y errors and X_{2B} can detect single Z or Y errors. In essence, these protocols measure a stabilizer of the Bell state (either ZZ or XX) and discard upon encountering an error. The only difference between Bell state distillation compared to other forms of error detection in Clifford circuits is that the stabilizer measurement happens in a distributed fashion; i.e., instead of computing the parity of the Bell state into

one qubit, we compute it into an entangled resource. This allows the parity checks to happen locally on each half of the Bell pair without further (quantum) communication across different halves of a Bell pair.

We can gain some intuition for the performance of the protocol by applying Z_{2B} to Bell states with a stochastic bit flip,

$$\rho_{02} = \mathcal{I} \otimes \mathcal{D}_p(\phi_{02}) \quad \text{and} \quad \rho_{13} = \mathcal{I} \otimes \mathcal{D}_q(\phi_{13}), \quad (1)$$

where ϕ is the maximally entangled state, the identity channel $\mathcal{I}$ acts on qubits 0 and 1, and the bit-flip channels $\mathcal{D}_p$ and $\mathcal{D}_q$ (see Appendix A) act on qubits 2 and 3, respectively, with bit-flip probabilities p and q , respectively. The measurements i and j in Fig. 1(b) equal each other with probability $p_s = (1-p)(1-q) + pq$ (sometimes called the acceptance probability or acceptance fraction) and result in a distilled state $\rho'_{02} = \mathcal{I} \otimes \mathcal{D}_r(\phi)$ where $r = pq/p_s$. The maximum Bell fidelity before distillation and Bell fidelity after distillation take values

$$F_b = \max(1-p, 1-q) \quad \text{and} \quad F_a = \frac{(1-p)(1-q)}{p_s}, \quad (2)$$

respectively. Interchanging p and q does not change F_a, F_b , and p_s , thus we let $p \leq q$, and write

$$F_a - F_b = \frac{1}{p_s} p(1-p)(1-2q) \geq 0, \quad (3)$$

where the inequality is strict whenever $0 < p \leq q < 1/2$. The strict inequality implies distillation improves Bell fidelity for all nontrivial p and q (under the assumption of a perfect implementation of the distillation protocol). This example demonstrates (1) recurrence in Fig. 1(a) reduces the X error rate from linear [$O(p)$ or $O(q)$] to quadratic, $O(pq)$, with probability p_s ; (2) the reduction occurs by postselecting away those errors on qubits 2 and 3 that anticommute with the Z_2Z_3 measurement. Similar conclusions hold for the X_{2B} recurrence protocol when analyzed for phase-flip noise, i.e., Z errors.

B. ZX_{3B} distillation

The limitation in the previous recurrence protocols was that they were not simultaneously sensitive to phase-flip and bit-flip error. If we increase the number of input Bell states to three (six qubits), then we can perform a distillation protocol with additional checks [17]. The circuit for the protocol is given in Fig. 1(c). One way to think about the protocol is to replace the ZZ measurements in recurrence with two sets of measurements, one set measuring ZZZ and another set measuring IXX (notice ZZZ and IXX commute with each other and thus can be measured simultaneously). Like recurrence, one accepts the final state only when measurement outcomes agree. A more general treatment of distillation protocols based on simultaneous measurement of commuting observable is available in Appendix C.

C. Distilling depolarized qubits

To better understand the performance of these distillation protocols for general noise, we look at these protocols when the Bell pairs are depolarized. In Fig. 1(a), suppose the Bell pairs are acted on by local depolarizing channels Λ (see

Appendix A for notation),

$$\rho_{02} = \mathcal{I} \otimes \Lambda_p(\rho) \quad \text{and} \quad \rho_{13} = \mathcal{I} \otimes \Lambda_q(\rho), \quad (4)$$

with probabilities p and q , respectively. Then the before and after distillation Bell fidelities are

$$F_b = \max(1-p, 1-q) \quad \text{and} \quad F_a = \frac{1}{p_s} \left((1-p)(1-q) + \frac{pq}{9} \right), \quad (5)$$

where

$$p_s = (1-2p/3)(1-2q/3) + 4pq/9 \quad (6)$$

is the acceptance probability (see Appendix D for details). In contrast to Eq. (3) where F_a is generally larger than F_b , here F_a is larger than F_b for a smaller set of initial Bell fidelities; in Fig. 2(b) we highlight this region where $F_a > F_b$ for roughly 0.19 fraction of the points sampled.

If we distill three Bell pairs (using the ZX_{3B} protocol) there is a comparatively larger region where the protocol shows improvements. In particular, assume the input to the circuit shown in Fig. 2(c) are Bell pairs acted on by local depolarizing channels,

$$\rho_{03} = \rho_{25} = \mathcal{I} \otimes \Lambda_p(\rho) \quad \text{and} \quad \rho_{14} = \mathcal{I} \otimes \Lambda_q(\rho), \quad (7)$$

where we assume two of the pairs have equal noise. Before and after distillation the Bell fidelities are

$$F_b = \max(1-p, 1-q) \quad \text{and} \quad F_a = \frac{1}{p_s} \left(p^2 \left(1 - \frac{28}{27}q \right) + p \left(\frac{19}{9}q - 2 \right) + 1 - q \right), \quad (8)$$

and the acceptance probability is

$$p_s = \frac{p^2}{9} \left(8 - \frac{32}{3}q \right) + \frac{p}{3} \left(\frac{20}{3}q - 5 \right) + 1 - q. \quad (9)$$

The set of initial Bell fidelities where $F_b > F_a$ is given in Fig. 2(c). In comparison, this region ($\simeq 0.61$ fraction of points) is about three times larger than the one in Fig. 2(b).

D. Global depolarizing noise

Next, we consider the effect of the protocols when the Bell pairs are subject to an n -qubit (global) depolarizing channel,

$$\mathcal{N}_\lambda(\rho) = (1-\lambda)\rho + \lambda \text{Tr}(\rho) \frac{\mathbb{I}_{2^n}}{2^n}. \quad (10)$$

The channel is simple to analyze, represents correlated noise among the qubits it acts on, and can be simulated in experiments (see Sec. IV A). We can use the formalism described in Appendix C to obtain an acceptance probability

$$p_G = (1-\lambda)p_a + \frac{\lambda}{2^{n-1}} \quad (11)$$

and Bell fidelity after distillation

$$F_G = \frac{1}{p_G} \left((1-\lambda)F p_a + \frac{\lambda}{2^{n+1}} \right), \quad (12)$$

where p_a and F are given in Eqs. (C1) and (C3), respectively. If $\rho_{AB} = \phi^{\otimes 2}$ and we use the Z_{2B} protocol (see Sec. II A) for distillation, then $p_a = 1$ in Eq. (C1), $F = 1$ in Eq. (C3),

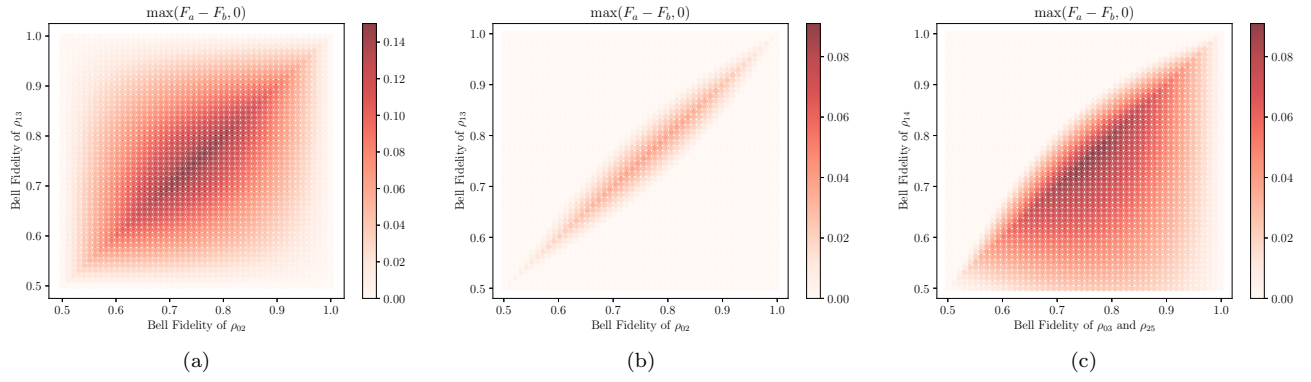


FIG. 2. Algebraically obtained increase in Bell fidelity by different distillation protocols. The x and y axes represent the Bell fidelity of the states prior to distillation, while the color corresponds to the increase (if any) in fidelity upon distillation. (a) Increase in Bell fidelity, see Eq. (3), using Z_{2B} distillation on Bell states undergoing bit flip errors (1), (b) Increase in Bell fidelity, see Eq. (5), using Z_{2B} distillation on Bell states undergoing depolarizing noise (4), and (c) Increase in Bell fidelity, see Eq. (8), using ZX_{3B} distillation on Bell states undergoing depolarizing noise (7).

fidelity of the state before distillation is $F_b = 1 - 3\lambda/4$, the acceptance probability and Bell fidelity after distillation take values

$$p_G = 1 - \frac{1}{2}\lambda \quad \text{and} \quad F_G = \frac{1}{p_G} \left(1 - \frac{7}{8}\lambda \right), \quad (13)$$

respectively. Notice $r := F_G/F_b > 1$ for all $0 < \lambda < 1$; i.e., distillation always improves the Bell fidelity for global depolarizing noise. This same conclusion holds when $\rho_{AB} = \phi^{\otimes 3}$ and the ZX_{3B} distillation protocol is applied. In this case,

$$p_G = 1 - \frac{3}{4}\lambda \quad \text{and} \quad F_G = \frac{1}{p_G} \left(1 - \frac{15}{16}\lambda \right). \quad (14)$$

III. CIRCUIT NOISE

Unlike the previous section, in this section we relax the assumption that the circuit used for distillation is perfect and includes gate noise and measurement error. The addition of these will greatly affect the ability to perform successful distillation. Motivated by the planar connectivity of superconducting qubits (and connecting to our experiments in the later sections), we consider circuits where Bell pairs need to be created locally and swapped so that the distillation circuit can be performed; see Fig. 3(a) as an example.

To model noise on measurements, we apply a bit-flip channel, $\mathcal{D}_m$, prior to measurement and noise on two-qubit gates is

modeled by adding a two-qubit (global) depolarizing channel, $\mathcal{N}_g$, on the two qubits involved in the gate. Other sources of noise, such as imperfections in initializing the qubits to $|0\rangle$ and those in implementing single-qubit gates, are ignored since these sources of noise can be comparatively smaller than measurement and two-qubit gate noise. In the following, we will consider distillation as a function of these noise parameters versus the fidelity of the input Bell pairs.

A. Local depolarizing noise

In the studies presented in this section, we vary the input Bell pair's fidelity using local depolarizing channels similar to Sec. II C. In Fig. 3 we show circuits for distilling Bell pairs under the noise model just described. Here, we add a qubit depolarizing channel in two places. These two places are before the first barrier (dotted vertical line) and between the second and third barriers in each of the subfigures of Fig. 3. Between each of these barriers, depolarizing noise is added to only one half of a Bell pair since noise on both halves of a Bell pair can be transferred to just one using the standard transpose trick [44]. This allows us to independently control the asymmetry among the Bell pairs and fidelity of the Bell pairs prior to distillation.

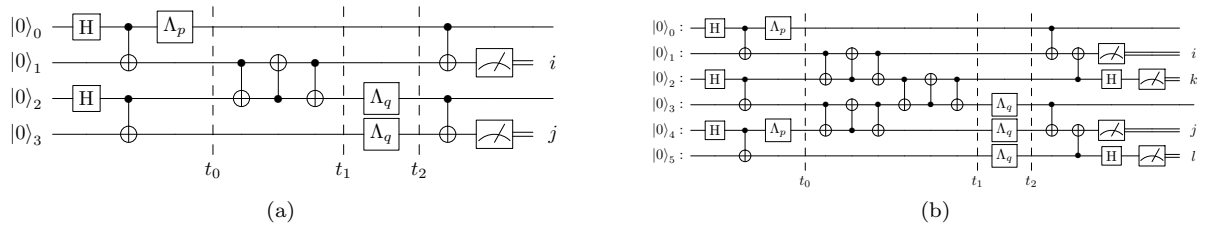
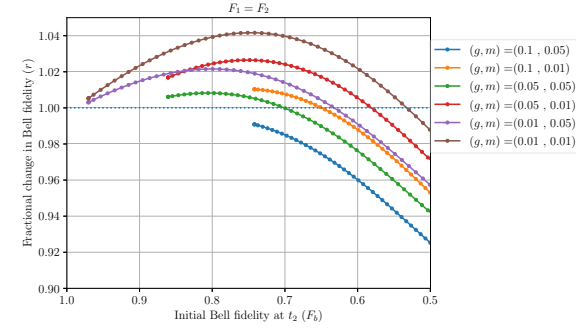
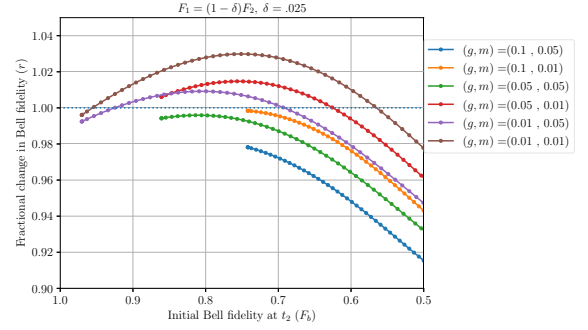


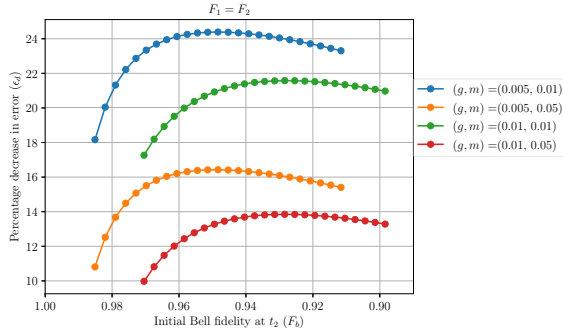
FIG. 3. Circuits for entanglement distillation with depolarizing channels Λ_p and Λ_q inserted at various stages. (a) Two local Bell pairs are created, swapped to form physically non-local Bell pairs, and then distilled using the Z_{2B} protocol (see Sec. III A 1 for additional details). and (b) Three local Bell pairs are created, swapped to form physically non-local Bell pairs, and then distilled using the ZX_{3B} distillation protocol (see Sec. III A 2 for additional details).



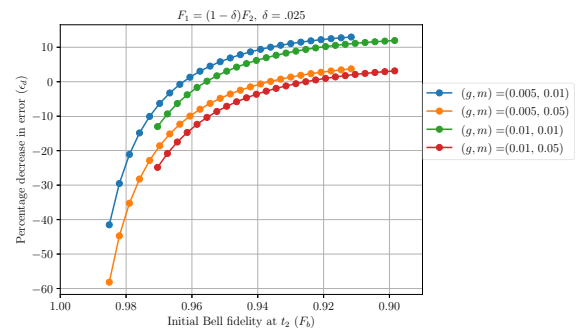
(a) Bell pairs are initially prepared with equal Bell fidelity, $F_1 = F_2$



(b) Bell pairs are initially prepared with unequal Bell fidelity, the first Bell pair has 2.5% lower Bell fidelity than the second.



(c) Bell pairs are initially prepared with equal Bell fidelity, $F_1 = F_2$



(d) Bell pairs are initially prepared with unequal Bell fidelity, the first Bell pair has 2.5% lower Bell fidelity than the second.

FIG. 4. Results from simulation of recurrence with circuit noise described in Fig. 3(a). Panels (a) and (b) show fractional change in Bell fidelity, r [see Eq. (15)], plotted against initial Bell fidelity F_b [defined below Eq. (15)], for various gate and measurement errors, g and m , respectively. The region where the plot remains above $r = 1$ indicates where the noisy distillation circuit is beneficial. Panels (c) and (d) focus on high-fidelity Bell pairs and show the percentage decrease in Bell infidelity, ϵ_d [defined in Eq. (16)], plotted against initial fidelity F_b for various gate and measurement errors.

1. Recurrence

We simulate the recurrence protocol (Z_{2B}) using the quantum circuit shown in Fig. 3(a). In this circuit, all four qubits are initialized to $|0\rangle$. First the circuit prepares two Bell states (across qubits 0–1 and 2–3) and then adds noise to the first Bell state (modeled via a depolarizing channel Λ_p acting on the top qubit). This noise helps model asymmetry in the initial Bell pairs. In the next part, after the first barrier (dotted line labeled t_0), the circuit swaps one half of each Bell pair creating a (physically) nonlocal Bell pair across qubits 0–2 and 1–3. Next, after the second barrier (dotted line labeled t_1), we apply a waiting error (modeled via a one-qubit depolarizing channel Λ_q acting on qubits 1 and 2 that constitute one half of each Bell pair). The final part of the circuit, after the third barrier, carries out a distillation protocol described in Fig. 1(a). All CNOT gates and measurements carry error, described by channels $\mathcal{N}_g$ and $\mathcal{D}_m$, respectively, as discussed in beginning of this section.

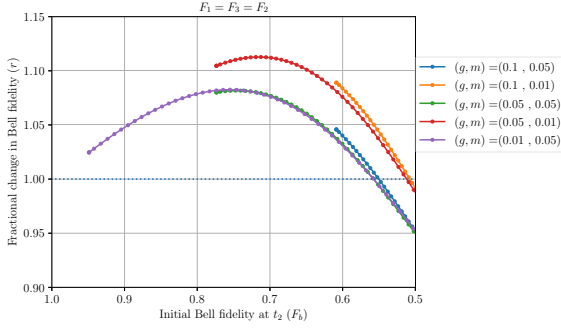
Variation in the performance of distillation with gate error g , measurement error m , and asymmetry between the initial Bell pairs is plotted in Fig. 4. In this figure, the ratio

$$r = F_a/F_b \quad (15)$$

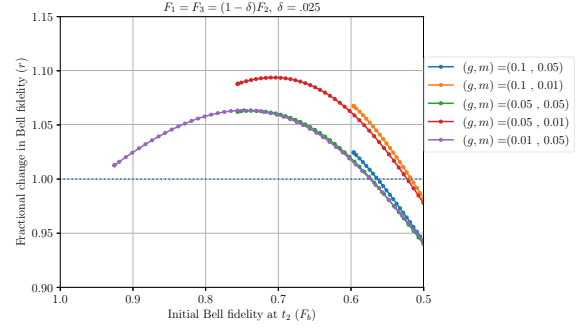
is plotted against F_b , where F_a is the Bell fidelity of the distilled Bell pair and F_b is the maximum of the Bell fidelities among the physically nonlocal Bell pairs just prior to being distilled. The Bell fidelities F_1 and F_2 are of the Bell pairs just prior to the first barrier in Fig. 3(a).

In the first plot [Fig 4(a)], the two Bell pairs have equal Bell fidelity; i.e., Λ_p in Fig. 3(a) is noiseless. For any fixed gate error $g \in \{0.01, 0.05, 0.1\}$ and measurement error $m \in \{0.01, 0.05\}$, we increase the parameter q in the waiting error Λ_q and plot r [Eq. (15)] as a function of the initial Bell fidelity F_b . In each of these plots as the waiting error q is increased from zero, the Bell fidelity prior to distillation, F_b , decreases, while the ratio r first increases, reaches a maximum, and then decreases. Across different plots, if we fix the gate error g but increase the measurement error m , we notice a decrease in r . On the other hand, if we fix the measurement error m , increasing the gate error g has two effects. First, it decreases the initial Bell fidelity F_b at $q = 0$ (i.e., the plot begins at a smaller F_b value) and also decreases the initial r value corresponding to that initial Bell fidelity. Second, it shrinks the interval of F_b values over which $r > 1$.

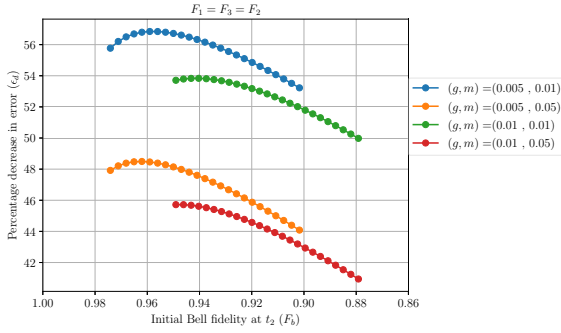
In the second plot [Fig 4(b)], the two Bell pairs have unequal Bell fidelity; i.e., Λ_p in Fig. 3(a) is noisy with $p \neq 0$. After the action of this noisy identity gate, the first Bell pair



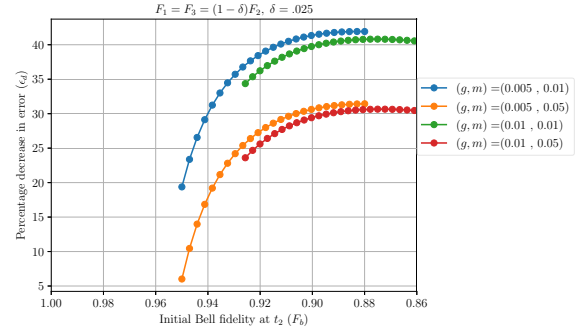
(a) Bell pairs are initially prepared with equal Bell fidelity, $F_1 = F_3 = F_2$



(b) Bell pairs are initially prepared with unequal Bell fidelity, the first and third Bell pair has 2.5% lower Bell fidelity than the second



(c) Bell pairs are initially prepared with equal Bell fidelity, $F_1 = F_3 = F_2$



(d) Bell pairs are initially prepared with unequal Bell fidelity, the first and third Bell pairs have 2.5% lower Bell fidelity than the second.

FIG. 5. Results from simulation of ZX_{3B} distillation with circuit noise given in Fig. 3(b). Panels (a) and (b) show fractional change in Bell fidelity, r [see Eq. (15)], plotted against initial Bell fidelity F_b [defined below Eq. (15)], for various gate and measurement errors. The region of the plot above $r = 1$ indicates where the noisy distillation circuit is beneficial. Panels (c) and (d) focus on low-noise Bell pairs and show the percentage decrease in Bell infidelity, ϵ_d [defined in Eq. (16)], plotted against initial fidelity F_b for various gate and measurement errors.

(across qubits 0 and 1) becomes more noisy than the second (across qubits 2 and 3) and Bell fidelity of the first Bell pair, F_1 , is 2.5% lower than that of the second, F_2 . Salient features of the plot in Fig. 4(b) resemble those of Fig. 4(a) discussed above. The key difference is that for fixed g and m values, the values of r in Fig. 4(b) are smaller than the corresponding values in Fig. 4(a). In fact, we only see improvement in a small region where the gate and measurement errors are $\leq 1\%$.

In Figs. 4(c) and 4(d), we focus on smaller gate and measurement errors, $g = 1 \times 10^{-2}$ or 5×10^{-3} and $m = 5 \times 10^{-2}$ or 1×10^{-2} . Here, the y axis is the percentage decrease in error, i.e.,

$$\epsilon_d = \frac{F_a - F_b}{1 - F_b} \times 100, \quad (16)$$

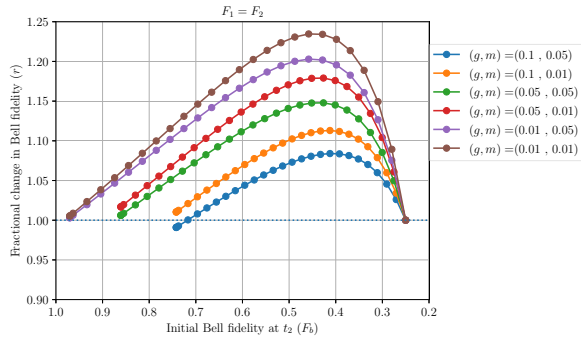
while the x axis is the initial Bell fidelity F_b . We obtain points along these axes by varying the waiting error q in Λ_q . Here, the initial Bell fidelity is typically greater than 0.9 and we notice a 10%–20% decrease in error by performing distillation on Bell pairs with equal Bell fidelity [see Fig. 4(c)]. When the Bell pairs have unequal Bell fidelity, recurrence does not necessarily decrease the error for very high Bell fidelity but manages to decrease error as the initial Bell fidelities become smaller [see Fig. 4(d)]. This observation is consistent with results in Fig. 2(a) that show that the range of initial Bell

fidelities where distillation improves is narrow when the initial Bell fidelities are high.

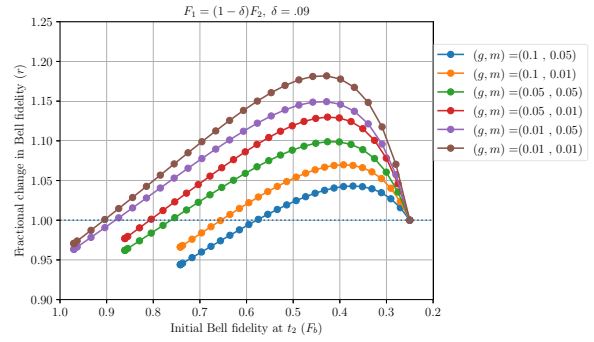
2. Three Bell distillation

Next, we simulate the ZX_{3B} protocol using the circuit in Fig. 3(b). This circuit prepares three Bell pairs, which we label as AA , BB , and CC , initialized on qubits 0–1, 2–3, and 4–5, respectively. Next, the circuit adds noise to Bell states AA and CC via Λ_p . After the first barrier (labeled t_0), a sequence of CNOT gates reorder the Bell pairs into $ABCABC$, such that one half of each Bell pair is separated from its other half by two qubits. After the second barrier, the circuit applies a waiting error (via channel Λ_q acting on one half of each Bell pair), and finally, it carries out the ZX_{3B} distillation protocol (see Sec. II B).

The first plot in Fig. 5(a) describes the variation in the ratio r as a function of the F_b for fixed gate and measurement errors. All three Bell pairs have the same initial Bell fidelity. The variation in r and F_b with the gate error g and measurement error m are similar to those in Fig. 4(a), described below Eq. (15). For fixed g and m , the ratio r in Fig. 5(a) is typically higher than those in Fig. 4(a) when $r > 1$. In addition, the value of F_b at which r shifts from a value greater than one to a value less than one is typically lower in Fig. 5(a) compared to Fig. 4(a); i.e., the parameter region and amount by which



(a) Bell pairs are prepared with equal Bell Fidelity.



(b) Bell pairs are prepared with unequal Bell fidelity, the second Bell pair has 9% lower Bell fidelity than the second.

FIG. 6. Results from simulation of recurrence Z_{2B} with global depolarizing noise [for circuit, see Fig. 11(a) in Appendix G]. Panels (a) and (b) show fractional change in Bell fidelity, r defined in Eq. (15), plotted against initial Bell fidelity F_b [defined below Eq. (15)], for various gate and measurement errors. The region of the plot above $r = 1$ indicates where the noisy distillation circuit is beneficial.

distillation provides an improvement seem to be typically larger in Fig. 5(a) compared to Fig. 4(a).

In Fig. 5(b), two of the Bell pairs, the first AA pair across qubits 0–1 and third CC across qubits 4–5, have lower Bell fidelity, than the second pair BB across qubits 2–3; i.e. $F_1 = F_3 < F_2$. The noise parameter p is such that the initial Bell fidelity $F_1 = F_3 = 0.975F_2$. As a result, distillation protocol here can be seen as an attempt to improve the fidelity of one Bell pair using two Bell pairs with lower fidelity. Variation in improvement r and F_b with the gate error g and measurement error m is similar to those in Fig. 4(b), described below Eq. (15). For fixed g and m , the ratio r in Fig. 5(a) can be higher than those in Fig. 4(b); however, this need not be the case in general, even when $r > 1$.

The focus in Figs. 5(c) and 5(d) is on smaller gate and measurement errors, $g = 1 \times 10^{-2}$ or 5×10^{-3} and $m = 5 \times 10^{-2}$ or 1×10^{-2} . As the error parameter q is varied, we plot the percentage decrease in error against the initial Bell fidelity F_b . When the Bell pairs being distilled have equal fidelity initially, we find a 40%–50% decrease in error, as shown in Fig. 5(c). The decrease in error is lower when the Bell pairs have unequal Bell fidelities initially, as shown in Fig. 5(d); it is possible that error increases if the initial Bell fidelities are made even more unequal.

B. Global depolarizing noise

In this section, we apply circuit noise to Bell pairs and degrade them under a global depolarizing noise channel $\mathcal{N}_\lambda$ (see Fig. 11 in Appendix G); this global noise is consistent with Clifford twirling [45,46], which we will do experimentally in Sec. IV A.

1. Two Bell distillation

We numerically study the performance of the recurrence distillation protocol (Z_{2B}) after the Bell pairs are degraded by a global depolarizing channel. The distillation circuit under study modifies the one in Fig. 3(a) by replacing local with global depolarizing noise during waiting [circuit available as Fig. 11(a) in Appendix G]. In the first plot [see Fig. 6(a)], the Bell pairs are created with equal Bell fidelity, $F_1 = F_2$. As the global depolarizing noise parameter λ is increased, the

fidelity of Bell pairs prior to distillation decreases. For any fixed gate and measurement error, as λ is increased, the ratio r , plotted on the y axis, first increases, reaches a maximum, and then decreases to one. This ratio can be as high as 1.2, and it remains above one for a wider range of gate and measurement error values. This represents broad improvement from distillation. The improvements shrink as the gate and measurement error are increased. In Fig. 6(b), the Bell pairs have unequal Bell fidelity prior to distillation. This asymmetry is created by noise channel Λ_p on qubit labeled 0 in Fig. 11(a). A key qualitative effect of this asymmetry is in the ratio r . This ratio remains below one for higher initial Bell fidelities, thus shrinking the range of gate and measurement error values over which distillation presents an improvement. The improvement is also smaller in comparison to those reported in Fig. 6(a).

2. Three Bell distillation

Here, we continue from the previous section, but indicate results from the double selection (ZX_{3B}) protocol. The effect of global depolarizing noise on the ZX_{3B} distillation protocol is similar to that of recurrence shown in Fig. 6 and described in Sec. III B 1. However, the value of r is typically higher than those in the corresponding plots in Fig. 6 (see Fig. 15 in Appendix H for plots).

IV. EXPERIMENTAL RESULTS AND DEVICE NOISE

Given the simulations of the previous sections, we now want to consider distillation protocols on a real device. We run experiments on a 127-qubit device with fixed-frequency qubits and fixed coupling, an IBM Eagle device `ibm_kyiv`. For the purposes of these experiments, we consider smaller sections of the device with linear connectivity and run each experiment in parallel across different sections of the device in order to obtain more statistics on noise. For more device details, see Appendix E.

For each circuit, we estimate the fidelity of the Bell state using direct fidelity estimation by measuring three circuits (see Appendix B for more details). However, this scheme does have measurement error. In contrast to Ref. [40], here we choose to not correct for measurement errors due to issues that can cause potentially nonphysical and or unexpectedly high

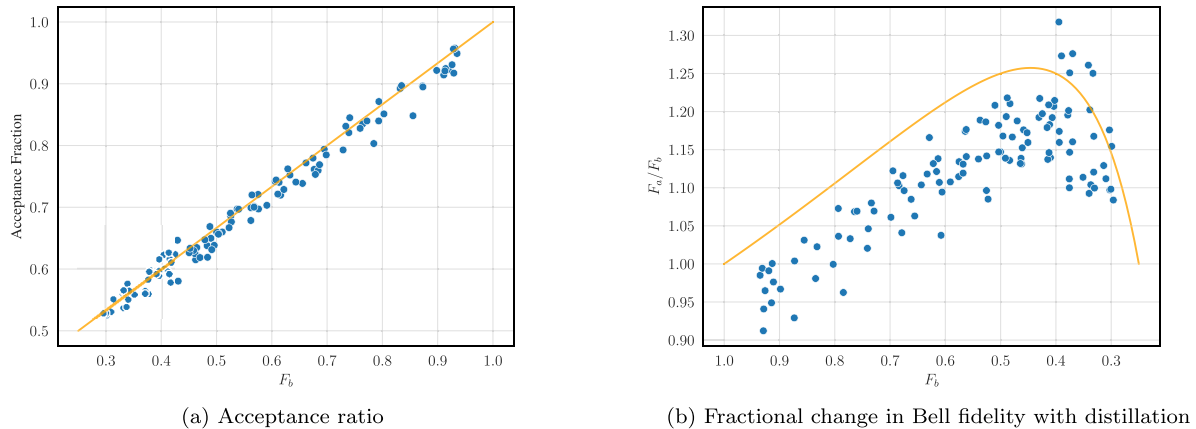


FIG. 7. Experimental results (blue) on `ibm_kyiv` for the Z_{2B} recurrence distillation protocol with global depolarizing noise. Orange curves are theory assuming perfect distillation.

fidelities (see, e.g., discussion in Ref. [47]). Our experimental results thus provide a good lower bound on the Bell fidelities.

A. Two Bell distillation experiment with global depolarizing noise

Following from the discussion on Sec. III B, first, we experimentally study the recurrence protocol under global depolarizing noise. This global depolarizing channel $\mathcal{N}_\lambda$ is implemented using layers of two-qubit Clifford circuits. The first half of such a circuit is composed of multiple applications of a length- k random sequence of two-qubit Clifford gates and the second half is the inverse of the first half. This inverse is simply the mirror image of the first half and these layers are called mirror Clifford layers [48] (see Fig. 12 in Appendix G for an example of such a circuit with two layers in the first half).

These measurements are carried out on a set of four adjacent qubits on a device. On 13 such sets we carry out the experiment outlined above for different fixed $k \in \{0, 2, 4, 6, 8, 10, 12\}$. For each k , we obtain an average of 13 sets of the before and after distillation Bell fidelities along with the acceptance probability for distillation. These are plotted in Fig. 7. In Fig. 7(a), we plot the acceptance fraction as a function of the initial Bell fidelity. The dots represent points from the experiment, while the straight line is obtained from numerics. In Fig. 7(b), we plot the ratio r as a function of the initial Bell fidelity prior to global depolarizing noise. The blue dots represent data, while the orange curve represents the theoretically expected curve. This theoretical curve from Eq. (13) assumes noiseless gates and measurement. There is fairly good agreement between the theory curves and the experiment data, particularly considering the x axis is subject to systematic error due to measurement error and the theory calculation assumes the distillation circuit is perfect. This demonstrates that under some conditions on the device we can see an improvement due to distillation; however, the improvement arises from a systematic procedure to degrade the Bell pairs via global depolarizing noise. Furthermore, high-fidelity Bell pairs do not improve in this experiment.

B. Two Bell distillation experiment with idle noise

Here, we consider a more natural experiment, create a Bell pair and wait a time t for the Bell fidelity to decay under device T_1/T_2 noise and then measure (note we apply a standard echo sequence to remove low frequency noise). These T_1/T_2 noise channels are not simply Pauli channels and they are not an effective error obtained from averaging over random iterations, as we did in the previous section. However, an effective T_1/T_2 idling error is not obtained by simply waiting on an actual device; due to ZZ interactions between neighboring pairs, the Bell state fidelity can strongly degrade. We show a simple example of this in Fig. 8. Here, we can get tremendous improvement in the Bell fidelity, but this is because we allowed the Bell fidelity to decrease via a coherent error term; the purity of the state is unchanged. One has to be careful of similar situations where distillation provides an apparent benefit, but where improvement could more easily be achieved without postselection. In this example, the ZZ coherent error term can also be canceled by applying staggered dynamical decoupling (DD) [49–51] [see Fig. 13(a) in Appendix G]. With staggered DD, the circuit is now well described by T_1 and T_2 errors.

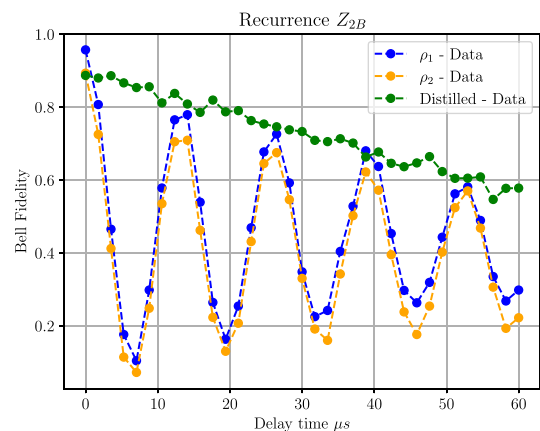
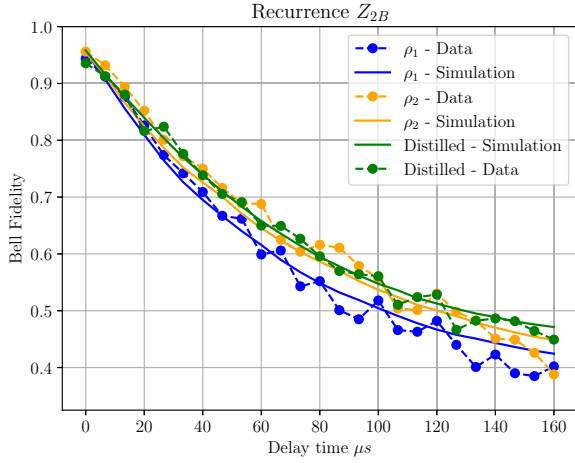
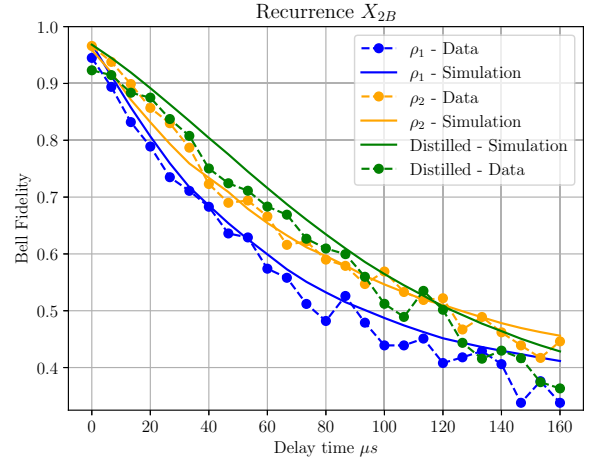


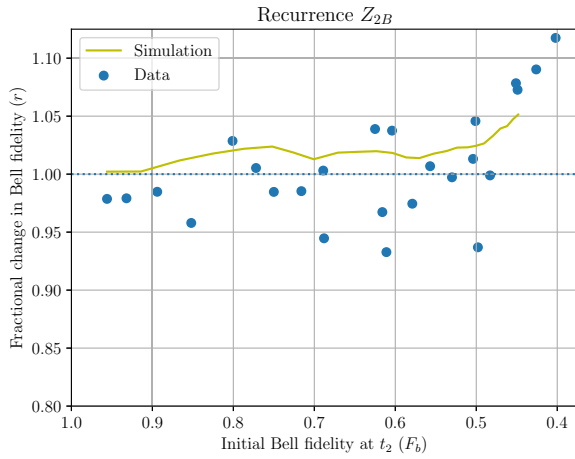
FIG. 8. Experimental data from qubits (0,1,2,3) of `ibm_kyiv` (points) showing increase in the Bell fidelity under Z_{2B} distillation of qubits undergoing coherent ZZ errors.



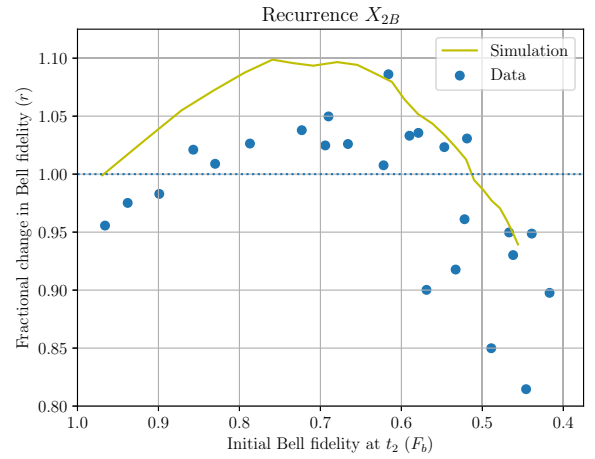
(a) Bell fidelity of the physically non-local Bell pairs prior to and after recurrence Z_{2B} .



(b) Bell fidelity of the physically non-local Bell pairs prior to and after recurrence X_{2B} .



(c) Fractional change in Bell fidelity with recurrence Z_{2B} .



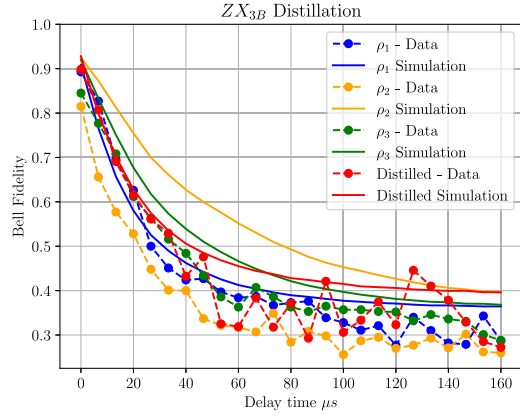
(d) Fractional change in Bell fidelity with recurrence X_{2B} .

FIG. 9. Data and simulation for recurrence Z_{2B} and X_{2B} on qubits (0,1,2,3) of `ibm_kyiv`. In panels (a) and (b), the Bell fidelity for various states is plotted against the delay time. States ρ_1 and ρ_2 label the different physically nonlocal Bell pairs being distilled, while the state obtained after distillation appears separately. Solid lines are simulations assuming reported noise parameters as described in Tables III and IV of Appendix E, while dash-dotted lines represent experimental data. Panels (c) and (d) plot the fractional change r defined in Eq. (15), against the initial Bell fidelity F_b defined below Eq. (15). Additional plots are shown in Figs. 16 and 17 in Appendix H.

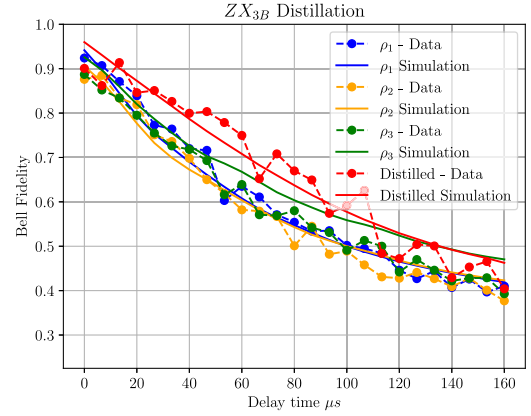
Once we apply staggered DD, typical data are shown in Fig. 9 with dotted lines. We see some natural features of the experiments. One, the starting Bell fidelity is lower than one due to the creation and swap operations [for instance, see the Bell fidelity in Figs. 9(a) and 9(b) at $t = 0$ is less than one]. Second, there is a natural asymmetry in the initial fidelity due to the variability of noise parameters between qubits [for instance, the Bell fidelity of the two bell pairs in Fig. 9(b) is unequal]. Third, depending on the distillation protocol used, Z_{2B} or X_{2B} , one may or may not see broad improvement in Bell fidelity [see Fig. 9(d) with improvement and Fig. 9(c) without]; however, we do not yet see broad across the device. To understand these features better, we look to build a more involved device noise model.

C. Device model and numerics

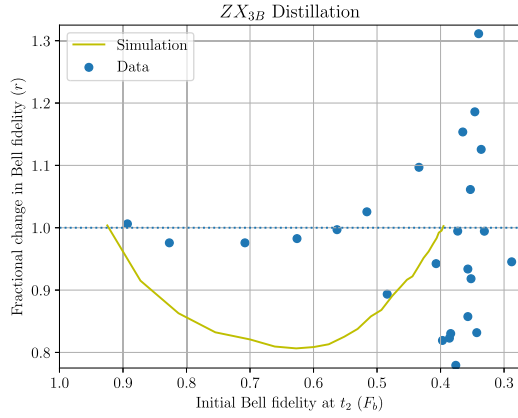
To build a more involved device model, we must include T_1/T_2 noise [42,43] (see Appendix F for a mathematical description) during the wait periods. These include when the swap is occurring, when the Bell states are idle, and when the measurements are occurring. Furthermore, due to the effect described in Ref. [52], it is not enough to use the bare T_1/T_2 even though we performed staggered echo sequences. We must still include the ZZ terms and simulate the echo sequence to get the proper decay of the Bell states during the wait. We continue to add-in the gate and measurement noise as described in Sec. III A. We pull the ZZ, T_1 , T_2 , gate times, and measurement error from the backend of the device (see Tables III and IV in Appendix E).



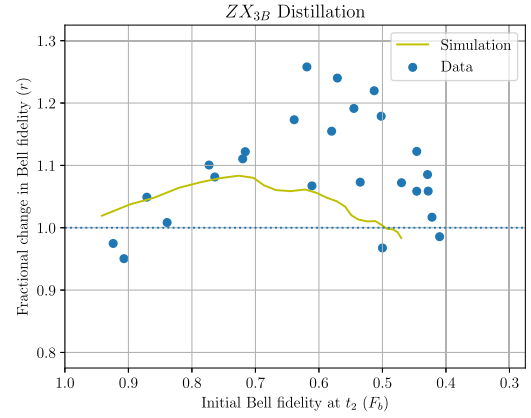
(a) Bell fidelity of the physically non-local Bell pairs prior to and after distillation on qubits (59,60,61,62,63,64). Simulations do not fit the data well; one possibility is that the coherence varied between the distillation run and when the coherence was measured.



(b) Bell fidelity of the physically non-local Bell pairs prior to and after distillation on qubits (3,4,5,6,7,8). Simulations fit the data well.



(c) Fractional change in Bell fidelity on qubits (59,60,61,62,63,64)



(d) Fractional change in Bell fidelity on qubits (3,4,5,6,7,8)

FIG. 10. Data and simulation for ZX_{3B} recurrence on `ibm_kyiv`. In panels (a) and (b), the Bell fidelity for various states is plotted against the delay time. States ρ_1 , ρ_2 , and ρ_3 label the different physically nonlocal Bell pairs being distilled, while the state obtained after distillation appears separately. Solid lines are simulations assuming reported noise parameters as described in Tables V and VI of Appendix E, while dash-dotted lines represent experimental data. Panels (c) and (d) plot the fractional change r defined in Eq. (15), against the initial Bell fidelity F_b defined below Eq. (15). Additional plots are shown in Figs. 18 and 19 in Appendix H.

For recurrence Z_{2B} , the simulations are given by the solid lines in Fig. 9(a) and we notice fairly good agreement between the experimental and numerical simulations. Confirming the general trend of the data is consistent with the calibration data. However, the figure does not indicate consistent improvement in Bell fidelity with distillation. To get improvements, we notice (see Table III in Appendix E) that, in general, the dominant source of noise on the qubits under consideration is T_2 noise (i.e., Z errors), while the Z_{B2} catches T_1 noise, i.e., X and amplitude damping-type errors.

The X_{2B} distillation protocol catches Z errors. Results from implementing this protocol in both simulation and experiment are given in Fig. 9(b). In this figure, there is reasonable agreement between the experimental data and numerical simulations. In addition, notice that there is general improvement in Bell fidelity, indicating that the likely dominant source of T_2 errors are caught by X_{2B} recurrence. Improvements are not seen for higher Bell fidelities. This likely occurs due to both measurement errors and asymmetry in Bell fidelities of the

two Bell pairs being distilled. The former are not removed by distillation, while the latter generally shrink the noise region where distillation shows improvement. In specific cases, simulations indicate that lowering measurement errors to zero can allow distillation to give improvements even for high-fidelity Bell pairs.

D. Three Bell distillation experiment with idle noise

Using the same technique for experiment and numbers as in the previous sections, we extend the experiment to the three Bell (ZX_{3B}) protocol (circuit available in Fig. 14 in Appendix G) and results of this are shown in Fig. 10. Again, we get reasonable agreement with the numerics using just the backend calibration parameters. We see two cases, in the first [see Fig. 10(a)] we have small acceptance probability and no real improvement in the Bell state; here, the noise parameters are quite large. In the second set of data [see Fig. 10(d)], we see some strong improvement from distillation, but this

does not persist well to the highest fidelities and therefore limits us from distributing higher-fidelity entanglement this way. In places where we find improvement [say at, e.g., 50 μ s in Fig. 10(b)], one can experience a loss of data, i.e., an acceptance probability below 50%.

V. CONCLUSIONS

We carried out both a numerical and an experimental analysis of some basic distillation protocols. The simplest numerical analysis in Sec. II adds asymmetry among the Bell pairs since the asymmetry can be inherent in the way multiple Bell pairs are arranged prior to distillation. That analysis points at regions in the noise parameter space where entanglement distillations can provide improvements. A more involved analysis using device characterization data, non-Pauli noise channels, and modeling the echo sequence in our experiment allows us to obtain reasonable agreement with the experimental data obtained from entanglement distillation. The noise models indicate how practical difficulties in distillation come from (1) a natural asymmetry among the Bell pairs being distilled; (2) understanding the noise which distillation must remove; (3) measurement noise; and (4) fixing a metric of success for distillation. When the metric of success is improving average Bell fidelity over several pairs or removing coherent errors in a single pair, we find that the simplest distillation protocols can provide broad improvements (see discussions in Secs. IV A and IV B). However, for using distilled Bell pairs in modular computing, it is valuable to consider a stricter metric: distill (physically) nonlocal Bell pair with higher fidelity than what can be obtained using other available means. This work leaves open the question of broadly improving upon this metric (see related theoretical work [32] in this direction). One way to obtain such improvements using the protocols discussed here would be to make noise more uniform on the Bell pairs; another would be to lower the measurement errors on the qubits; while a third would be to go beyond the types of protocols discussed in this work.

In our analysis, we found it particularly useful to view distillation as stabilizer checks. This made it “easy” to view different distillation protocols as tools to catch different types of errors (which need not be Pauli errors, and are not necessarily modeled as such here either). This can help guide the selection of distillation protocols depending on the source of the dominant noise in one’s error model.

A variety of prior work focuses on the optimal trade-offs between success probability and Bell fidelity upon distillation over Pauli noise channels. For practical studies, our work motivates the study of T_1/T_2 noise channels (these have experimentally available noise parameters) and benchmarking protocols when starting with unequal bell pairs that incur measurement errors.

ACKNOWLEDGMENTS

We thank Sam Stein, Michael DeMarco, Andrew Cross, and John Smolin for helpful discussions. This work was supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers,

Co-design Center for Quantum Advantage (C2QA) under Contract No. DE-SC0012704.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

APPENDIX A: NOTATION

Let $\{|0\rangle, |1\rangle\}$ denote the standard basis of a qubit Hilbert space $\mathcal{H}$ (sometimes called the Z basis), $|\pm\rangle := (|0\rangle \pm |1\rangle)/\sqrt{2}$ denote the so-called Hadamard basis, and $H := |+\rangle\langle 0| + |-\rangle\langle 1|$ represent the Hadamard gate. Let $X = |0\rangle\langle 1| + |1\rangle\langle 0|$, $Y = i|0\rangle\langle 1| - i|1\rangle\langle 0|$, and $Z = |0\rangle\langle 0| - |1\rangle\langle 1|$, together with the 2×2 identity, $\mathbb{I}$, denote the Pauli matrices $\mathcal{P}$. Any two square matrices P, Q are said to commute when $PQ = QP$ and anticommute when $PQ = -QP$. Note that any two Pauli matrices either commute or anticommute.

The qubit X -dephasing (bit-flip) channel,

$$\mathcal{D}_q(\rho) = (1 - q)\rho + qX\rho X, \quad (\text{A1})$$

applies the Pauli X operator with probability $0 \leq q \leq 1/2$. The qubit depolarizing channel,

$$\begin{aligned} \Lambda_p(\rho) &= (1 - p)\rho + \frac{p}{3}(X\rho X + Y\rho Y + Z\rho Z), \\ &= \left(1 - \frac{4}{3}p\right)\rho + \frac{4p}{3}\text{Tr}(\rho)\frac{\mathbb{I}}{2}, \end{aligned} \quad (\text{A2})$$

applies Pauli X, Y , and Z operators with equal probability $p/3$ and the identity operator $\mathbb{I}$ with probability $1 - p$ where $0 \leq p \leq 1$. We use $\mathcal{I}$ to denote the identity channel, $\mathcal{I}(\rho) = \rho$ for all operators ρ .

The fidelity of a two-qubit density operator ρ with the maximally entangled state,

$$|\phi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}, \quad (\text{A3})$$

is what we call the Bell fidelity, $F := \langle\phi|\rho|\phi\rangle$. The Bell fidelity can be estimated directly from Pauli measurements (see Appendix B). We refer to $1 - F$ as the infidelity with the Bell state (A3). The two-qubit CNOT gate with 0th qubit as control and first qubit as target is $\text{CNOT}_{01} = |0\rangle\langle 0| \otimes \mathbb{I} + |1\rangle\langle 1| \otimes X$.

The tensor product of Pauli matrices $\mathcal{P}$ on n qubits is sometimes called a Pauli string. We use a compressed notation to represent such a string by suppressing the identity matrix and using subscript to denote the system label. For instance, with $n = 3$, we denote $Z \otimes \mathbb{I} \otimes X$ by Z_0X_2 .

APPENDIX B: DIRECT BELL FIDELITY ESTIMATION

A two-qubit state ρ has Bell fidelity,

$$F = \langle\phi|\rho|\phi\rangle = \frac{1}{4}(1 + \langle ZZ\rangle_\rho + \langle XX\rangle_\rho - \langle YY\rangle_\rho), \quad (\text{B1})$$

where $\langle N\rangle_\rho := \text{Tr}(N\rho)$. The Bell fidelity can be computed by measuring expectation value of ZZ, XX , and YY operators. To compute $\langle ZZ\rangle_\rho$, one may measure each qubit of ρ in the Z basis, obtain probabilities p_{ij} for measurement outcome corresponding to basis $|ij\rangle$, and then evaluate $\langle ZZ\rangle_\rho = 2(p_{00} + p_{11}) - 1$.

Expectation values for XX and YY can be computed in an analogous manner by measuring qubits in the X and Y basis instead of Z . Thus, to measure the Bell fidelity this way, one does three different measurement experiments, one for each Pauli measurement basis. Given a circuit for measuring a qubit in the Z basis, one can measure in the X and Y basis by applying H and $H \cdot S^\dagger$, respectively, prior to measuring in the Z basis, where $S = \text{diag}(1, -i)$ is $\sqrt{Z}$.

APPENDIX C: A GENERAL PROTOCOL FOR DISTILLATION

Both recurrence and the ZX_{3B} protocols in Sec. II can be viewed as special cases of those in Refs. [53,54]. These distillation protocols can be viewed in a unified way using a slightly different notation as follows:

Let $\mathcal{H}_{Ai}$ and $\mathcal{H}_{Bi}$ each be a qubit Hilbert space, $\mathcal{H}_i := \mathcal{H}_{Ai} \otimes \mathcal{H}_{Bi}$ be a two-qubit Hilbert space, $\mathcal{H}_A := \bigotimes_{i=1}^n \mathcal{H}_{Ai}$ and $\mathcal{H}_B := \bigotimes_{i=1}^n \mathcal{H}_{Bi}$ each be n -qubit Hilbert spaces, and ρ_{AB} be a $2n$ -qubit state. To distill a single two-qubit state from ρ_{AB} , one applies some unitary $U = U_A \otimes U_B$ to ρ_{AB} then postselects for agreement on Z measurements made on $2(n-1)$ qubits. Let the unmeasured system be labeled 0. To describe the postselected state, consider $T = \mathbb{I}_0 \otimes L$, where $\mathbb{I}_0$ is the identity on $\mathcal{H}_0$, $L = \bigotimes_{i=1}^{n-1} P_i$, and $P_i = |00\rangle\langle 00| + |11\rangle\langle 11|$ is a projector defined on $\mathcal{H}_i$. The postselection succeeds with probability

$$p_a = \text{Tr}(TU \rho U^\dagger T^\dagger) \quad (\text{C1})$$

and results in a state

$$\rho' = \frac{1}{p_a} \text{Tr}_{i \neq 0}(TU \rho U^\dagger T^\dagger), \quad (\text{C2})$$

where the partial trace is over all spaces $\mathcal{H}_i$ except $i = 0$. The fidelity of the postselected state is

$$F = \text{Tr}(\rho' \phi). \quad (\text{C3})$$

APPENDIX D: DISTILLED FIDELITY CALCULATION

As explained at the end of Sec. II, the recurrence protocol postselects away errors on qubits 2 and 3 that anticommute with $Z_2 Z_3$. This idea can be used to compute the final fidelity F_a in Eq. (5) and acceptance probability p_s in Eq. (6). In Fig. 1(a), suppose

$$\rho_{02} = \mathcal{I} \otimes \mathcal{P}_p(\rho) \quad \text{and} \quad \rho_{13} = \mathcal{I} \otimes \mathcal{P}_q(\rho), \quad (\text{D1})$$

where $\mathcal{P}_p(\rho) := p_I \rho + p_x X \rho X + p_y Y \rho Y + p_z Z \rho Z$ is a qubit Pauli channel, each $p_i \geq 0$, and $\sum_{i \in \{I, x, y, z\}} p_i = 1$. The collection of the Pauli errors not caught by recurrence are detailed in Table I. In this table, the first column denotes errors on qubits 2 and 3 that commute with $Z_2 Z_3$, and are thus not postselected away by recurrence, the second column denotes the probability of these errors, and the third column denotes the effective error after postselection on qubit labeled 2. Setting $p_x = p_y = p_z = p/3$ and $q_x = q_y = q_z = q/3$ and summing the entries in the second column gives the acceptance probability p_s [Eq. (6)], while summing those entries in the second column where the effective error is I_2 and normalizing by p_s gives the Bell fidelity of the distilled state F_a [Eq. (5)].

For the ZX_{3B} distillation protocol, we may derive the acceptance probability p_s in Eq. (9) and Bell fidelity upon

TABLE I. Accepted errors by recurrence and their effect: Pauli error not postselected away by recurrence (see Sec. II A), their probability, and their effect on the final distilled Bell pair.

Error	Probability	Transformed error
$\mathbb{I}$	$p_I q_I$	$\mathbb{I}_2$
Z_3	$p_I q_z$	Z_2
$X_2 X_3$	$p_x q_x$	X_2
$X_2 Y_3$	$p_x q_y$	Y_2
$Y_2 X_3$	$p_y q_x$	Y_2
$Y_2 Y_3$	$p_y q_y$	X_2
Z_2	$p_z q_I$	Z_2
$Z_2 Z_3$	$p_z q_z$	$\mathbb{I}_2$

acceptance F_b in Eq. (8) using a procedure analogous to the one above. Suppose in Fig. 2(c),

$$\begin{aligned} \rho_{03} &= \mathcal{I} \otimes \mathcal{P}_p(\rho), \\ \rho_{14} &= \mathcal{I} \otimes \mathcal{P}_q(\rho), \quad \text{and} \\ \rho_{25} &= \mathcal{I} \otimes \mathcal{P}_r(\rho), \end{aligned} \quad (\text{D2})$$

then we list the collection of errors not caught by the ZX_{3B} distillation protocol in Table II. In this table, the first, second, and third columns represent the accepted errors (those which commute with both $Z_3 Z_4 Z_5$ and $X_4 X_5$), their probability and their effect on qubit 3 after qubits 4 and 5 are measured, respectively. Setting $p_x = p_y = p_z = r_x = r_y = r_z = p/3$, and $q_x = q_y = q_z = q/3$ and summing the entries in the second column gives the acceptance probability p_s in Eq. (9), while summing the entries in the second column corresponding to no error ($\mathbb{I}$) in the third and normalizing by p_s gives the fidelity after distillation F_a in Eq. (8).

TABLE II. Accepted errors by ZX_{3B} and their effect: Pauli error not postselected away by the ZX_{3B} protocol (see Sec. II B), their probability, and their effect on the final distilled Bell pair.

Error	Probability	Transformed error
$\mathbb{I}$	$p_I q_I r_I$	$\mathbb{I}_3$
$X_4 X_5$	$p_I q_x r_x$	$\mathbb{I}_3$
$Z_4 Z_5$	$p_I q_z r_z$	Z_3
$Y_4 Y_5$	$p_I q_y r_y$	Z_3
$X_3 X_5$	$p_x q_I r_x$	X_3
$X_3 X_4$	$p_x q_x r_I$	X_3
$X_3 Z_4 Y_5$	$p_x q_z r_y$	Y_3
$X_3 Y_4 Z_5$	$p_x q_y r_z$	Y_3
Z_3	$p_z q_I r_I$	Z_3
$Z_3 X_4 X_5$	$p_z q_x r_x$	Z_3
$Z_3 Z_4 Z_5$	$p_z q_z r_z$	$\mathbb{I}_3$
$Z_3 Y_4 Y_5$	$p_z q_y r_y$	$\mathbb{I}_3$
$Y_3 X_5$	$p_y q_I r_x$	Y_3
$Y_3 X_4$	$p_y q_x r_I$	Y_3
$Y_3 Z_4 Y_5$	$p_y q_z r_y$	X_3
$Y_3 Y_4 Z_5$	$p_y q_y r_z$	X_3

TABLE III. Single-qubit and two-qubit error rates from `ibm_kyiv` last updated on 14 May 2024 at 07:45:06 UTC used for simulation of recurrence Z_{2B} . Measurement delay is 1.24 μs . The CNOT error rate is taken to be the echoed cross-resonance (ECR) error rate and kept the same when role of target and control qubits is reversed.

Qubit	T_1 (μs)	T_2 (μs)	Measurement error
0	257.944	323.573	6.5×10^{-3}
1	477.815	224.595	9.1×10^{-3}
2	263.123	123.047	4.3×10^{-3}
3	260.839	232.639	4.6×10^{-3}
Qubit 1	Qubit 2	ZZ rate (Hz)	ECR error
0	1	-52 860.4	$7.751\,53 \times 10^{-3}$
1	2	-55 319.3	10.3203×10^{-3}
2	3	-45 908	4.2953×10^{-3}

APPENDIX E: DEVICE OVERVIEW

Experiments were carried out on IBM's fixed-frequency transmon superconducting processor, `ibm_kyiv`. This is a 127-qubit chip with qubits arranged in a heavy-hexagonal lattice which reduces crosstalk with reasonable overhead in circuit layout mapping. This processor is from IBM's Eagle processor family that uses the echoed cross-resonance gate for its entangling gate and features multiplexed readout. To support the higher qubit count, the chip features multilayer wiring with care taken to reduce the effects of quantum and classical crosstalk. This architecture results in a median T_1 and T_2 of 276 and 122 μs , respectively, and median ECR gate error and readout error of 1.1×10^{-2} and 6.0×10^{-3} , respectively. Tables III–VI give error parameters for specific qubits on which distillation experiments and simulation were discussed in the main text.

APPENDIX F: T_1/T_2 CHANNEL

Noise on a superconducting qubit can be described using a damping-dephasing channel [42] that may be expressed

TABLE IV. Single-qubit and two-qubit error rates from `ibm_kyiv` last updated on 15 May 2024 at 16:05:02 UTC used for simulation of recurrence X_{2B} . Measurement delay is 1.24 μs . The CNOT error rate is taken to be the ECR error rate and kept the same when role of target and control qubits is reversed.

Qubit	T_1 (μs)	T_2 (μs)	Measurement error
0	276.892	312.245	2.5×10^{-3}
1	512.747	218.116	2.6×10^{-3}
2	236.636	98.102	4.2×10^{-3}
3	330.719	232.639	11.2×10^{-3}
Qubit 1	Qubit 2	ZZ rate (Hz)	ECR error
0	1	-52 860.4	$4.434\,72 \times 10^{-3}$
1	2	-55 319.3	$8.103\,92 \times 10^{-3}$
2	3	-45 908	$4.037\,14 \times 10^{-3}$

TABLE V. Single-qubit error rates from `ibm_kyiv` last updated on 17 May 2024 at 05:54:35 UTC used for simulation of the three bell experiment. Measurement delay is 1.24 μs .

Qubit	T_1 (μs)	T_2 (μs)	Measurement error
3	410.738	232.639	4.2×10^{-3}
4	429.253	152.675	16.9×10^{-3}
5	365.064	384.404	6.8×10^{-3}
6	294.473	146.729	1.6×10^{-3}
7	339.754	371.239	3.4×10^{-3}
8	458.343	259.075	2.7×10^{-3}
59	269.232	78.7512	2.4×10^{-3}
60	273.893	285.543	6×10^{-3}
61	318.205	152.633	7.9×10^{-3}
62	255.428	25.5405	23.6×10^{-3}
63	275.333	115.334	7.3×10^{-3}
64	231.769	47.6543	3.8×10^{-3}

as [43]

$$\mathcal{M}(\rho) = \sum_i O_i \rho O_i^\dagger, \quad (\text{F1})$$

where $O_0 = \sqrt{1-p}|0\rangle\langle 0| + \sqrt{1-g}|1\rangle\langle 1|$, $O_1 = \sqrt{g}|0\rangle\langle 1|$, $O_2 = \sqrt{p}|0\rangle\langle 0| - \sqrt{1-g}|1\rangle\langle 1|$, $0 \leq p \leq 1/2$ represents dephasing, and $0 \leq g \leq 1$ represents damping probability. The channel maps an input density operator with Bloch vector $\mathbf{r} = (x, y, z)$ to an output with Bloch coordinates $((1-2p)\sqrt{1-g}x, (1-2p)\sqrt{1-g}y, (1-g)z+g)$. When the output coordinates are parametrized as $(e^{-t/T_2}x, e^{-t/T_2}y, e^{-t/T_1}z + 1 - e^{-t/T_1})$, then

$$g = 1 - e^{-t/T_1}, \quad p = (1 - e^{-(1/T_2 - 1/2T_1)t})/2, \quad (\text{F2})$$

and $2T_1 \geq T_2$. When $p = 0$ or $T_2 = 2T_1$, $\mathcal{M}$ is an amplitude damping channel mapping $|1\rangle$ to $|0\rangle$ with probability g . When $g = 0$ or $T_1 = \infty$, $\mathcal{M}$ is a pure dephasing channel that applies a Z error with probability p . If a qubit with fixed T_1 and T_2 parameters waits idle for a time t , then channel modeling noise on this qubit, $\mathcal{M}_t$, is the $\mathcal{M}$ channel with parameters g and p depending on t , T_1 , and T_2 , as indicated in Eq. (F2).

TABLE VI. Two-qubit error rate and ZZ rate from `ibm_kyiv` last updated on 17 May 2024 at 05:54:35 UTC used for simulation of the three bell distillation protocol. The CNOT error rate is taken to be the same when role of target and control qubits is reversed.

Qubit 1	Qubit 2	ZZ rate (Hz)	CNOT error
3	4	-48 982	$8.498\,46 \times 10^{-3}$
4	5	-39 813.1	11.4476×10^{-3}
5	6	-76 347.7	$8.780\,52 \times 10^{-3}$
6	7	-57 303.5	16.352×10^{-3}
7	8	-40 264.1	$8.809\,76 \times 10^{-3}$
59	60	-12 7831	13.0765×10^{-3}
60	61	-38 618.7	$7.858\,57 \times 10^{-3}$
61	62	-57 210.3	13.1797×10^{-3}
62	63	-55 771.1	13.9869×10^{-3}
63	64	-40 636.1	$6.077\,38 \times 10^{-3}$

APPENDIX G: CIRCUITS FOR NOISE AND DISTILLATION

Here, we present the circuits used in the distillation protocols of the main text, as well as the modifications to those circuits that enabled us to study them under various types of noise.

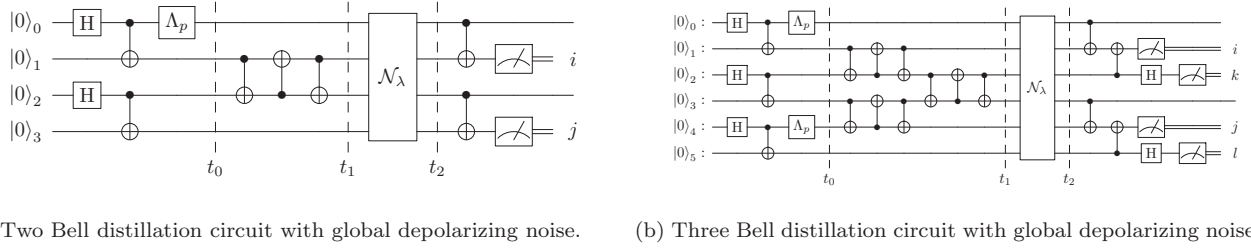


FIG. 11. Circuits for simulation of entanglement distillation with global depolarizing channel $\mathcal{N}_\lambda$ replacing the local depolarizing noise between t_1 and t_2 barriers in Fig. 3.

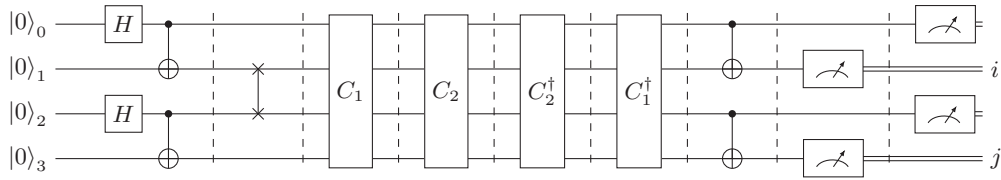
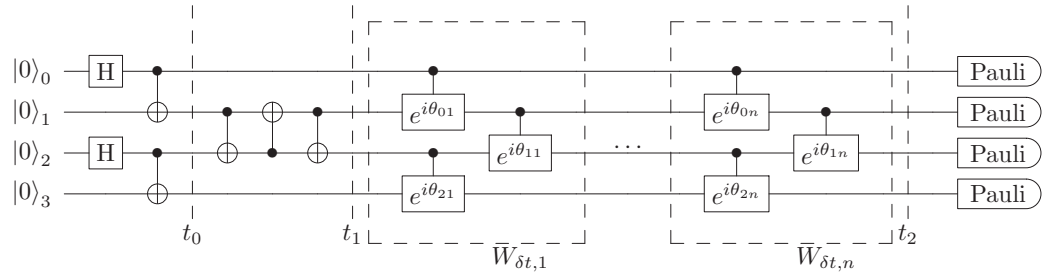


FIG. 12. Circuit for implementing recurrence under global depolarizing noise using mirror Clifford layers. Bell pairs are created before the first barrier (dotted vertical line), then qubits 1 and 2 are swapped to create physically nonlocal Bell pairs, next two layers of Clifford circuits C_1, C_2 and their inverse are applied, and followed by a recurrence protocol for distillation and measurement of all qubits.



(a) Two bell creation with idling noise.

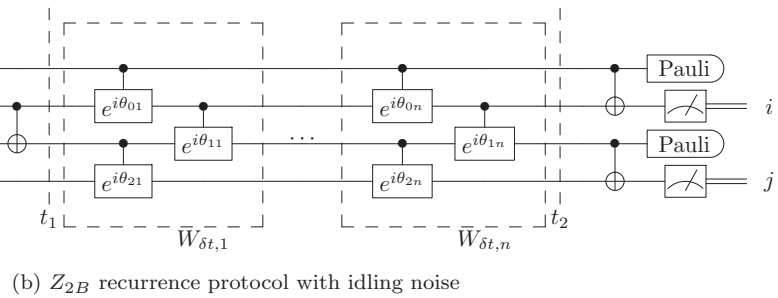
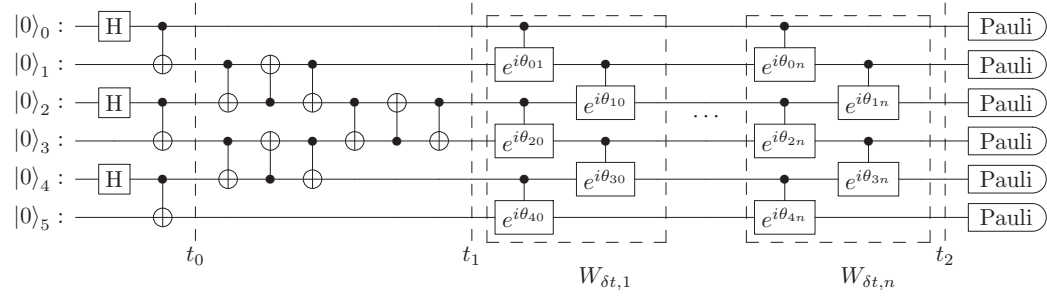


FIG. 13. Circuit(s) for the Z_{2B} recurrence protocol with idling noise. In both circuits, two Bell pairs are created prior to time label t_0 , between t_0 and t_1 the first and second qubits are swapped and a delay is added from t_1 to t_2 . The delay interval $t_2 - t_1$ is parametrized as $n\delta t$ where δt is variable and fixed integer n is the number of times the ZZ sequence $\bar{W}_{\delta t,k}$ is applied. Control phase angles in each sequence are $\theta_{jk} = f_k 2\pi \omega_j \delta t / n$ where $1 \leq k \leq n$, $0 \leq j \leq 2$, $f_k = 1$ if $k < n/2$ and -1 otherwise, and ω_j is the ZZ frequency between qubits j and $j + 1$. In panel (a) after time t_2 , Pauli basis measurements are carried out on each qubit to determine the Bell fidelity (see Appendix B for details). In panel (b) after time t_2 , ZZ measurements are carried out on qubits 1 and 3, while qubits 0 and 2 are measured in the Pauli basis for Bell fidelity estimation.



(a) Three bell creation with idling noise.

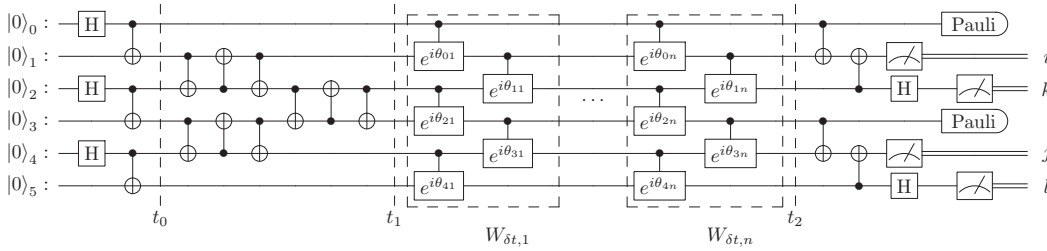
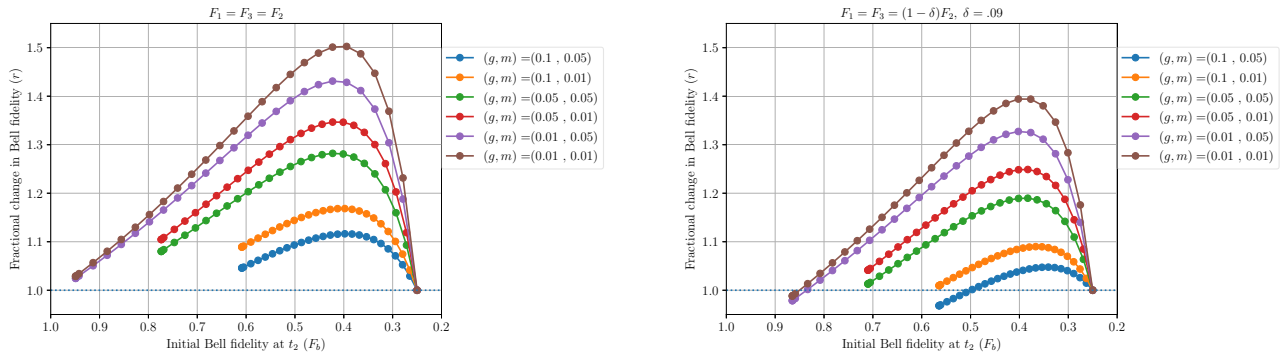
(b) ZX_{3B} distillation with idling noise

FIG. 14. Circuit(s) for ZX_{3B} distillation with idling noise. In both circuits, three Bell pairs are created prior to time label t_0 , between t_0 and t_1 the Bell pairs are swapped (see Sec. III A 2 for details) and a delay is added from t_1 to t_2 (in a manner analogous to that in Fig. 13). In panel (a) after time t_2 , Pauli basis measurements are carried out on each qubit to determine the Bell fidelity (see Appendix B for details). In panel (b) after time t_2 , ZZZ and IXX measurements are carried out (as explained in Sec. II B), while qubits 0 and 3 are measured in the Pauli basis for Bell fidelity estimation.

APPENDIX H: ADDITIONAL PLOTS FROM SIMULATION AND EXPERIMENTS

Here, we present plots generated using both simulations and experiments of entanglement distillation protocols discussed in the main text. Plots from simulation extend the ones in the main text, and those with experimental data separate out the Bell fidelity already presented and also show the acceptance probability.



(a) Bell pairs are prepared with equal Bell Fidelity.

(b) Bell pairs are prepared with unequal Bell fidelity, the first and third have the same Bell fidelity which is 9% less than that of the second Bell pair.

FIG. 15. Results from simulation of recurrence ZX_{3B} with global depolarizing noise [for circuit, see Fig. 11(b) in Appendix G]. Panels (a) and (b) show fractional change in Bell fidelity, r , defined in Eq. (15), plotted against initial Bell fidelity F_b [defined below Eq. (15)], for various gate and measurement errors. The region of the plot above $r = 1$ indicates where the noisy distillation circuit is beneficial.

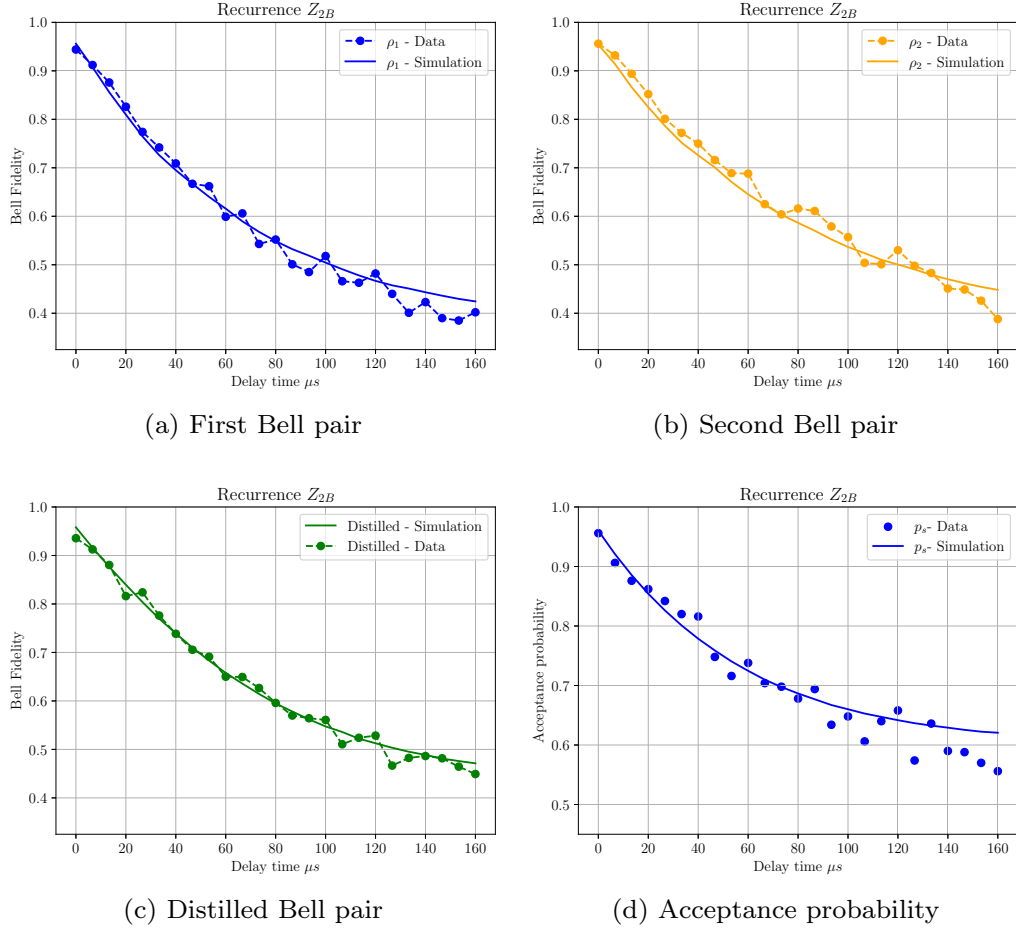


FIG. 16. Z_{2B} recurrence distillation on qubits [0,1,2,3] of `ibm_kyiv` (points). Here, simulations fit the data well. Solid lines are simulations assuming reported noise parameters as described in Table III of Appendix E. Panels (a), (b), and (c) plot Bell fidelity against delay time for the respective physically nonlocal state labeled in the figure, and panel (d) shows the acceptance probability vs delay time.

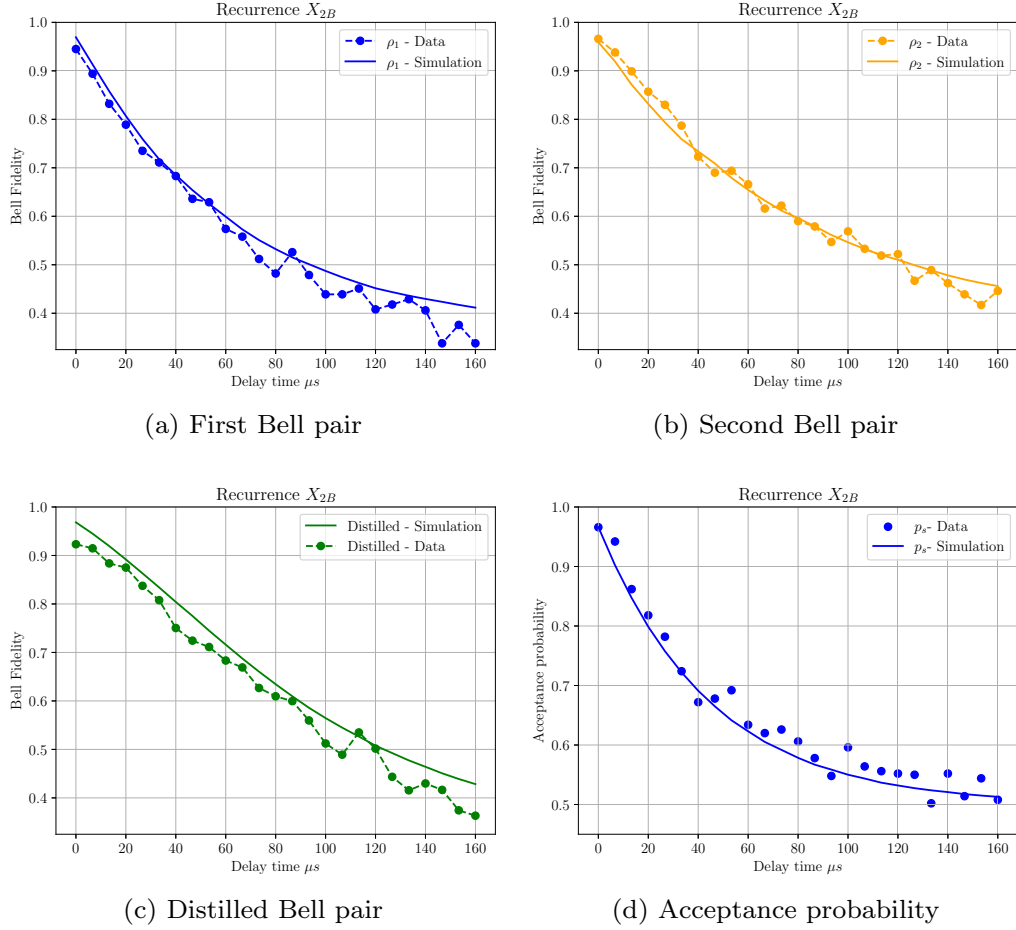


FIG. 17. X_{2B} recurrence distillation on qubits [0,1,2,3] of `ibm_kyiv` (points). Here, simulations fit the data well. Solid lines are simulations assuming reported noise parameters as described in Table IV of Appendix E. Panels (a), (b), and (c) plot Bell fidelity against delay time for the respective physically nonlocal state labeled in the figure, and panel (d) shows the acceptance probability vs delay time.

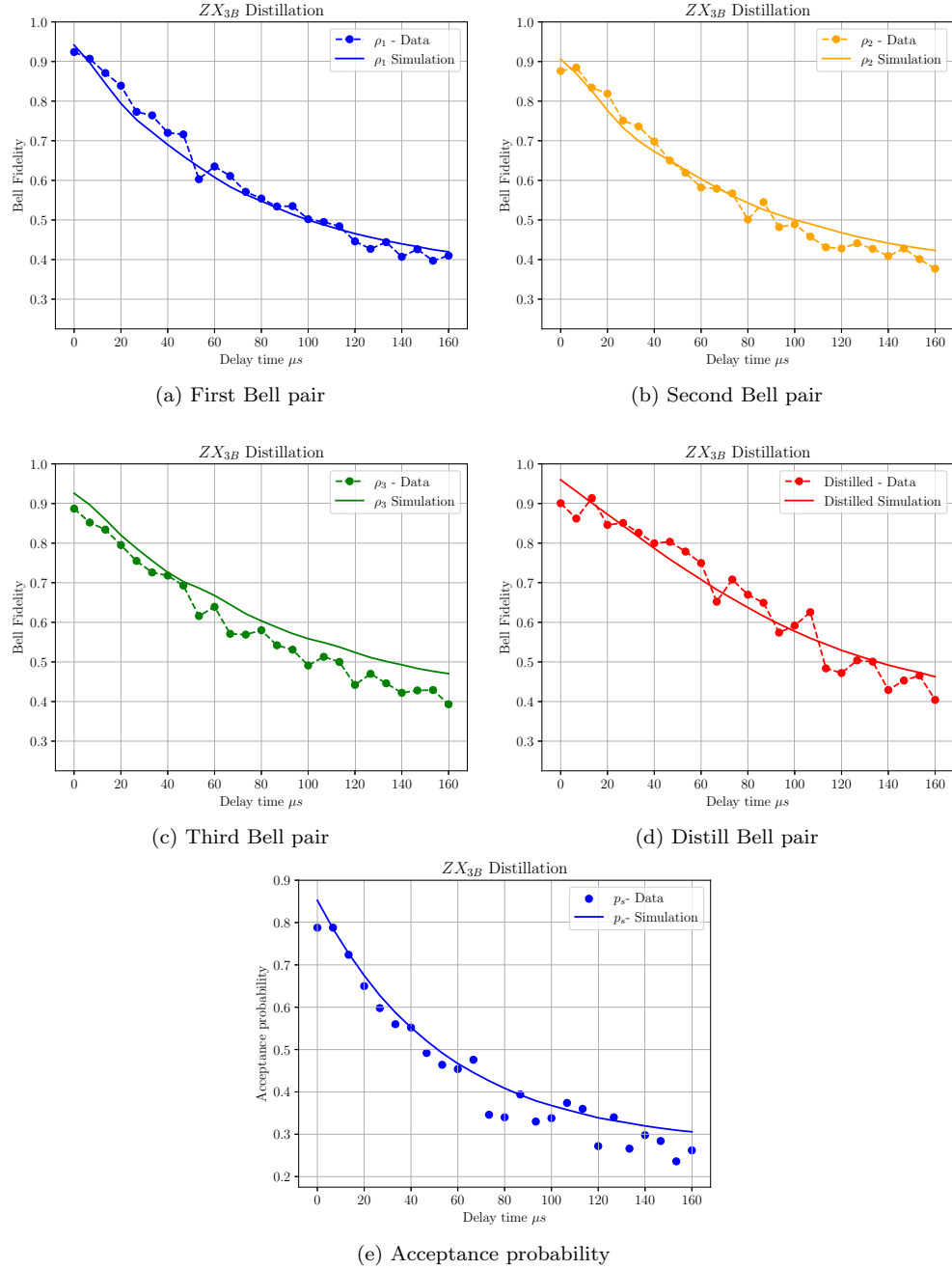


FIG. 18. Bell fidelity as a function of delay time for pairs before and after carrying out ZX_{3B} distillation on qubits [3, 4, 5, 6, 7, 8] of `ibm_kyiv` (points). Solid lines are simulations assuming reported noise parameters as described in Tables V and VI of Appendix E. Panels (a), (b), (c), and (d) plot Bell fidelity against delay time for the respective physically nonlocal state labeled in the figure, and panel (e) shows the acceptance probability vs delay time.

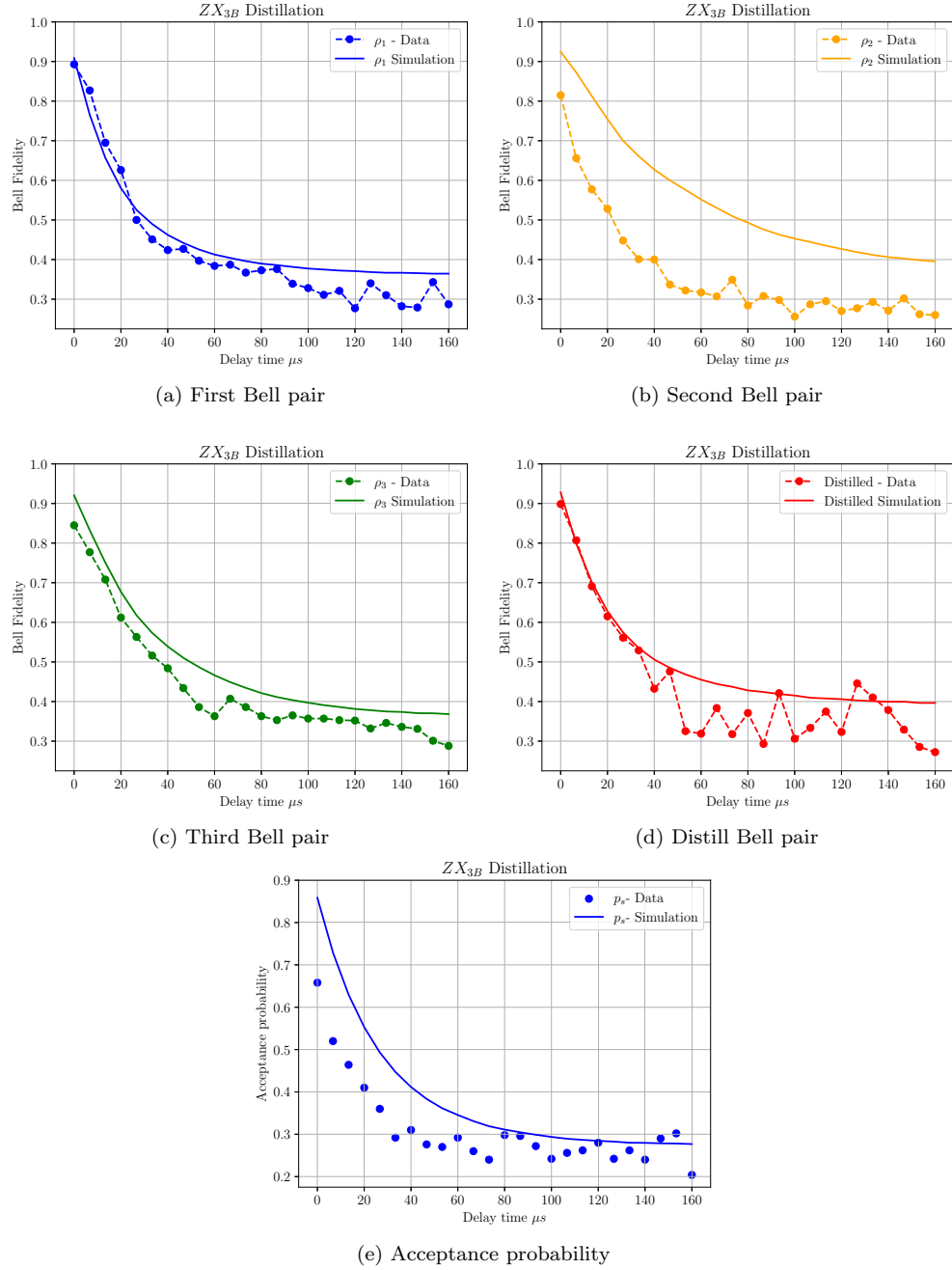


FIG. 19. Bell fidelity as a function of delay time for pairs before and after carrying out ZX_{3B} distillation on qubits [59, 60, 61, 62, 63, 64] of `ibm_kyiv` (points). Solid lines are simulations assuming reported noise parameters as described in Tables V and VI of Appendix E. Panels (a), (b), (c), and (d) plot Bell fidelity against delay time for the respective physically nonlocal state labeled in the figure, and panel (e) shows the acceptance probability vs delay time.

- [1] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Purification of noisy entanglement and faithful teleportation via noisy channels, *Phys. Rev. Lett.* **76**, 722 (1996).
- [2] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Mixed-state entanglement and quantum error correction, *Phys. Rev. A* **54**, 3824 (1996).
- [3] W. Dür, H.-J. Briegel, J. I. Cirac, and P. Zoller, Quantum repeaters based on entanglement purification, *Phys. Rev. A* **59**, 169 (1999).
- [4] C. Gidney, Tetrally compact entanglement purification, [arXiv:2311.10971](https://arxiv.org/abs/2311.10971).
- [5] P. Shor, Fault-tolerant quantum computation, in *Proceedings of 37th Conference on Foundations of Computer Science* (Burlington, VT, USA, IEEE, 1996), pp. 56–65.
- [6] E. Knill, Scalable quantum computing in the presence of large detected-error rates, *Phys. Rev. A* **71**, 042322 (2005).
- [7] S. Bravyi, G. Smith, and J. A. Smolin, Trading classical and quantum computational resources, *Phys. Rev. X* **6**, 021043 (2016).
- [8] J. Kim, J. Yun, and J. Bae, Purification of noisy measurements and faithful distillation of entanglement, *J. Phys. A Math. Theor.* **58**, 03LT01 (2024).
- [9] Y. Shi, C. Liu, S. Stein, M. Wang, M. Zheng, and A. Li, Design of an entanglement purification protocol selection module, [arXiv:2405.02555](https://arxiv.org/abs/2405.02555).
- [10] P. Yan, L. Zhou, W. Zhong, and Y. Sheng, Advances in quantum entanglement purification, *Sci. China Phys. Mech. Astron.* **66**, 250301 (2023).
- [11] A. Kang, S. Guha, N. Rengaswamy, and K. P. Seshadreesan, Trapped ion quantum repeaters with entanglement distillation based on quantum LDPC codes, in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)* (Bellevue, WA, USA, IEEE, 2023), Vol. 01, pp. 1165–1171.
- [12] J. Miguel-Ramiro, F. Riera-Sabat, and W. Dür, Quantum repeater for W states, *PRX Quantum* **4**, 040323 (2023).
- [13] L. Vandr  and O. G hne, Entanglement purification of hypergraph states, *Phys. Rev. A* **108**, 062417 (2023).
- [14] D. Devulapalli, E. Schoute, A. Bapat, A. M. Childs, and A. V. Gorshkov, Quantum routing with teleportation, *Phys. Rev. Res.* **6**, 033313 (2024).
- [15] S. Jansen, K. Goodenough, S. de Bone, D. Gijswijt, and D. Elkouss, Enumerating all bilocal Clifford distillation protocols through symmetry reduction, *Quantum* **6**, 715 (2022).
- [16] J. Miguel-Ramiro and W. D r, Efficient entanglement purification protocols for d -level systems, *Phys. Rev. A* **98**, 042309 (2018).
- [17] K. Fujii and K. Yamamoto, Entanglement purification with double selection, *Phys. Rev. A* **80**, 042308 (2009).
- [18] S. Glancy, E. Knill, and H. M. Vasconcelos, Entanglement purification of any stabilizer state, *Phys. Rev. A* **74**, 032319 (2006).
- [19] H. Aschauer and H. Briegel, Entanglement purification with noisy apparatus can be used to factor out an eavesdropper, *Eur. Phys. J. D* **18**, 171 (2002).
- [20] J.-W. Pan, C. Simon,  . Brukner, and A. Zeilinger, Entanglement purification for quantum communication, *Nature (London)* **410**, 1067 (2001).
- [21] C. A. Pattison, G. Baranes, J. P. B. Ataides, M. D. Lukin, and H. Zhou, Constant-rate entanglement distillation for fast quantum interconnects, in *Proceedings of the 52nd Annual International Symposium on Computer Architecture* (Association for Computing Machinery, New York, NY, 2025), pp. 257–270.
- [22] Y. Shi, A. Patil, and S. Guha, Stabilizer entanglement distillation and efficient fault-tolerant encoders, *PRX Quantum* **6**, 010339 (2025).
- [23] Y. Shi, A. Patil, and S. Guha, Measurement-based entanglement distillation and constant-rate quantum repeaters over arbitrary distances, *Phys. Rev. Lett.* **135**, 130803 (2025).
- [24] H.-J. Briegel, W. D r, J. I. Cirac, and P. Zoller, Quantum repeaters: The role of imperfect local operations in quantum communication, *Phys. Rev. Lett.* **81**, 5932 (1998).
- [25] X.-M. Hu, C.-X. Huang, Y.-B. Sheng, L. Zhou, B.-H. Liu, Y. Guo, C. Zhang, W.-B. Xing, Y.-F. Huang, C.-F. Li, and G.-C. Guo, Long-distance entanglement purification for quantum communication, *Phys. Rev. Lett.* **126**, 010503 (2021).
- [26] S. Stein, S. Sussman, T. Tomesh, C. Guinn, E. Tureci, S. F. Lin, W. Tang, J. Ang, S. Chakram, A. Li, M. Martonosi, F. T. Chong, A. A. Houck, I. L. Chuang, and M. A. DeMarco, HetArch: Heterogeneous microarchitectures for superconducting quantum systems, in *Proceedings of the 56th Annual IEEE/ACM International Symposium on Microarchitecture* (Association for Computing Machinery, New York, NY, 2023), pp. 539–554.
- [27] J. Ang, G. Carini, Y. Chen, I. Chuang, M. Demarco, S. Economou, A. Eickbusch, A. Faraon, K.-M. Fu, S. Girvin *et al.*, ARQUIN: Architectures for multinode superconducting quantum computers, *ACM Trans. Quantum Comput.* **5**, 59 (2024).
- [28] F. Rozpedek, T. Schiet, L. P. Thinh, D. Elkouss, A. C. Doherty, and S. Wehner, Optimizing practical entanglement distillation, *Phys. Rev. A* **97**, 062333 (2018).
- [29] S. Krastanov, V. V. Albert, and L. Jiang, Optimized entanglement purification, *Quantum* **3**, 123 (2019).
- [30] M. Vitoria, S. Tserkis, S. Krastanov, A. S. De La Cerda, S. Willis, and P. Narang, Entanglement purification on quantum networks, *Phys. Rev. Res.* **5**, 033171 (2023).
- [31] X. Zhao, B. Zhao, Z. Wang, Z. Song, and X. Wang, Practical distributed quantum information processing with LOCCNet, *npj Quantum Inf.* **7**, 159 (2021).
- [32] A. Zang, X. Chen, E. Chitambar, M. Suchara, and T. Zhong, No-go theorems for universal entanglement purification, *Phys. Rev. Lett.* **134**, 190803 (2025).
- [33] P. G. Kwiat, S. Barraza-Lopez, A. Stefanov, and N. Gisin, Experimental entanglement distillation and ‘hidden’ non-locality, *Nature (London)* **409**, 1014 (2001).
- [34] J.-W. Pan, S. Gasparoni, R. Ursin, G. Weihs, and A. Zeilinger, Experimental entanglement purification of arbitrary unknown states, *Nature (London)* **423**, 417 (2003).
- [35] T. Yamamoto, M. Koashi,  . K.  zdemir, and N. Imoto, Experimental extraction of an entangled photon pair from two identically decohered pairs, *Nature (London)* **421**, 343 (2003).
- [36] R. Dong, M. Lassen, J. Heersink, C. Marquardt, R. Filip, G. Leuchs, and U. L. Andersen, Experimental entanglement distillation of mesoscopic quantum states, *Nat. Phys.* **4**, 919 (2008).
- [37] L.-K. Chen, H.-L. Yong, P. Xu, X.-C. Yao, T. Xiang, Z.-D. Li, C. Liu, H. Lu, N.-L. Liu, L. Li, T. Yang, C.-Z. Peng, B. Zhao, Y.-A. Chen, and J.-W. Pan, Experimental nested purification

- for a linear optical quantum repeater, *Nat. Photonics* **11**, 695 (2017).
- [38] R. Reichle, D. Leibfried, E. Knill, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Experimental purification of two-atom entanglement, *Nature (London)* **443**, 838 (2006).
- [39] N. Kalb, A. A. Reiserer, P. C. Humphreys, J. J. W. Bakermans, S. J. Kamerling, N. H. Nickerson, S. C. Benjamin, D. J. Twitchen, M. Markham, and R. Hanson, Entanglement distillation between solid-state quantum network nodes, *Science* **356**, 928 (2017).
- [40] H. Yan, Y. Zhong, H.-S. Chang, A. Bienfait, M.-H. Chou, C. R. Conner, E. Dumur, J. Grebel, R. G. Povey, and A. N. Cleland, Entanglement purification and protection in a superconducting quantum network, *Phys. Rev. Lett.* **128**, 080504 (2022).
- [41] H. Qin, M.-M. Du, X.-Y. Li, W. Zhong, L. Zhou, and Y.-B. Sheng, Efficient multipartite entanglement purification with non-identical states, *arXiv:2311.10250*.
- [42] P. Aliferis, F. Brito, D. P. DiVincenzo, J. Preskill, M. Steffen, and B. M. Terhal, Fault-tolerant computing with biased-noise superconducting qubits: A case study, *New J. Phys.* **11**, 013061 (2009).
- [43] V. Siddhu, D. Abdelhadi, T. Jochym-O'Connor, and J. Smolin, Entanglement sharing across a damping-dephasing channel, in *2024 IEEE International Symposium on Information Theory (ISIT)* (Athens, Greece, IEEE, 2024), pp. 1432–1437.
- [44] The transposition trick transfers any single-qubit operator A acting on one half of a Bell pair $|\phi\rangle$ to the other half using the identity $\mathbb{I} \otimes A|\phi\rangle = A^T \otimes \mathbb{I}|\phi\rangle$.
- [45] J. Emerson, R. Alicki, and K. Życzkowski, Scalable noise estimation with random unitary operators, *J. Opt. B: Quantum Semiclassical Opt.* **7**, S347 (2005).
- [46] C. Dankert, R. Cleve, J. Emerson, and E. Livine, Exact and approximate unitary 2-designs and their application to fidelity estimation, *Phys. Rev. A* **80**, 012304 (2009).
- [47] R. S. Gupta, N. Sundaresan, T. Alexander, C. J. Wood, S. T. Merkel, M. B. Healy, M. Hillenbrand, T. Jochym-O'Connor, J. R. Wootton, T. J. Yoder, A. W. Cross, M. Takita, and B. J. Brown, Encoding a magic state with beyond break-even fidelity, *Nature (London)* **625**, 259 (2024).
- [48] T. Proctor, S. Seritan, K. Rudinger, E. Nielsen, R. Blume-Kohout, and K. Young, Scalable randomized benchmarking of quantum computers using mirror circuits, *Phys. Rev. Lett.* **129**, 150502 (2022).
- [49] Z. Zhou, R. Sitler, Y. Oda, K. Schultz, and G. Quiroz, Quantum crosstalk robust quantum control, *Phys. Rev. Lett.* **131**, 210802 (2023).
- [50] L. Shirizly, G. Misguich, and H. Landa, Dissipative dynamics of graph-state stabilizers with superconducting qubits, *Phys. Rev. Lett.* **132**, 010601 (2024).
- [51] A. Seif, H. Liao, V. Tripathi, K. Krsulich, M. Malekakhlagh, M. Amico, P. Jurcevic, and A. Javadi-Abhari, Suppressing correlated noise in quantum computers via context-aware compiling, in *2024 ACM/IEEE 51st Annual International Symposium on Computer Architecture (ISCA)* (Buenos Aires, Argentina, IEEE, 2024), pp. 310–324.
- [52] P. Jurcevic and L. C. G. Góvia, Effective qubit dephasing induced by spectator-qubit relaxation, *Quantum Sci. Technol.* **7**, 045033 (2022).
- [53] A. Ambainis and D. Gottesman, The minimum distance problem for two-way entanglement purification, *IEEE Trans. Inf. Theory* **52**, 748 (2006).
- [54] P. W. Shor and J. A. Smolin, Quantum error-correcting codes need not completely reveal the error syndrome, *arXiv:quant-ph/9604006*.